

**Univerzita Karlova**

**Filozofická fakulta**

Ústav románských studií

# **Bakalářská práce**

Alžběta Vítková

## **Le pronom *on* en français et ses équivalents en tchèque et en espagnol**

Francouzské zájmeno *on* a jeho protějšky v češtině a španělštině

French pronoun *on* and its counterparts in Czech and Spanish

Praha 2018

Vedoucí práce: PhDr. Olga Nádvorníková, Ph.D

**Prohlášení:**

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 25. 7. 2018

Alžběta Vítková

**Klíčová slova (česky)**

Francouzština – španělština – zájmeno *on* – překlad – InterCorp – vágní referent – determinace

**Klíčová slova (anglicky):**

French – Spanish – pronoun *on* – translation – InterCorp – vague pronoun reference – determination

## **Abstrakt**

Bakalářská práce *Le pronom **on** en français et ses équivalents en tchèque et en espagnol* se zabývá pluralitou referentů zájmena *on* ve francouzštině, a to zejména třemi stupni jejich determinace. Teoretická část představuje různé pohledy na toto zájmeno a jeho význam, naznačen je jeho historický vývoj i možnosti překladu do češtiny a španělštiny. Korpusový výzkum na korpusu InterCorp ukázal, že ačkoli ekvivalenty v těchto dvou jazycích ve většině případů nespádají do stejného formálního typu vyjádření, ve více než polovině případů se shodují ve stupni determinace. Z výzkumu vyplývá, vágní referent zájmena *on* vykazuje jistou závislost na typu textu a bezprostředním i širším kontextu, přesto jej nelze ve všech případech objektivně určit. Zároveň je patrná tendence právních textů užívat neosobní vyjádření, tedy nejnižší stupeň determinace, zatímco v beletristických textech jsou stupně determinace rozloženy téměř rovnoměrně.

## **Abstract**

This bachelor's thesis *French pronoun **on** and its counterparts in Czech and Spanish* deals with the referential plurality of the French pronoun *on*. A special attention is paid to the degree of reference's determination. The theoretical part presents different perspectives of the study of this pronoun and of its meaning. Then, evolution of its use as well as types of its counterparts in Czech and Spanish are introduced. The empirical part analyses data excerpted from the parallel corpus InterCorp. It demonstrates that counterparts in Czech and Spanish are usually classified in different formal types. However, the degrees of reference determination in Czech and Spanish correspond in more than a half of instances. Research shows that the reference's determination is defined by the context (immediate, large and extralinguistic as well), nevertheless an objective universal determination remains impossible. Moreover, legal texts tend to employ the lowest degree of determination while the degrees of determination in fiction seem to be quite equally spread.

## **Poděkování**

Na tomto místě bych ráda poděkovala PhDr. Olze Nádvorníkové, Ph.D. za vstřícný přístup a cenné rady a připomínky při zpracování této bakalářské práce. Za podporu bych chtěla poděkovat také své rodině.

## Table des matières

INTRODUCTION .....	8
1 PARTIE THÉORIQUE .....	10
1.1 Pronom <i>on</i> , défi à de nombreuses facettes .....	10
1.1.1 Pronom <i>on</i> du point de vue stylistique .....	11
1.1.2 Usages historiques du pronom <i>on</i> et l'insertion de « <i>l'</i> » .....	12
1.1.3 Appartenance du pronom <i>on</i> à la classe de mots .....	13
1.2 Détermination du référent du <i>on</i> .....	14
1.2.1 Valeurs référentielles du pronom <i>on</i> .....	14
1.2.2 Accord de l'adjectif ou du participe passé .....	16
1.3 Répertoire des équivalents du pronom <i>on</i> en tchèque et en espagnol .....	17
1.3.1 Equivalents du pronom <i>on</i> en espagnol .....	18
1.3.2 Equivalents du pronom <i>on</i> en tchèque .....	21
2 PARTIE MÉTHODOLOGIQUE .....	25
2.1 Présentation des corpus utilisés .....	25
2.2 Conditions de la recherche .....	26
2.3 Méthodologie de traitement des données .....	27
2.4 Ecueils rencontrés .....	28
2.5 Questions de recherche .....	29
3 PARTIE EMPIRIQUE .....	32
3.1 Pronom <i>on</i> du point de vue de fréquence et de son voisinage .....	32
3.1.1 Voisinage de la variante <i>l'on</i> .....	33
3.1.2 Accord du participe et de l'adjectif avec le sujet du verbe <i>être</i> .....	34
3.1.3 Types des verbes en voisinage du pronom <i>on</i> .....	35
3.2 Analyse des équivalents dans les textes littéraires .....	41
3.2.1 Correspondances entre les équivalents du pronom <i>on</i> en tchèque et en espagnol dans les textes littéraires .....	43
3.2.2 Echelle de la détermination dans les textes littéraires .....	45

3.3 Analyse des équivalents dans les textes juridiques.....	55
3.3.1 Correspondances entre les équivalents du pronom <i>on</i> en tchèque et en espagnol dans les textes juridiques .....	57
3.3.2 Echelle de la détermination dans les textes juridiques .....	59
3.4 Comparaison des échelles de la détermination .....	70
4 CONCLUSION.....	74
5 PISTES POUR DES RECHERCHES ULTÉRIEURES .....	78
Listes des tableaux et figures .....	79
Listes des exemples analysés .....	80
Références bibliographiques.....	85
Corpus consultés .....	87
Résumé en français .....	88
Résumé v češtině .....	90

# INTRODUCTION

Le pronom *on* est une des spécificités du système des pronoms personnels français. Il offre une vaste palette de référents : selon le contexte, il peut se référer à une personne précise ou à un ensemble précis. Or, c'est aussi un ensemble générique ou bien une vérité générale qui peuvent être dénotés. Enfin, le référent de ce même pronom peut être totalement effacé, indéterminé.

Ainsi, il est évident qu'il s'agit d'un phénomène assez complexe. Les équivalents directs du pronom *on* se trouvent dans d'autres langues, telles que l'allemand (*man*). Cependant, comme l'indique Tláskal (2005, p. 377), dans les langues comme le tchèque et l'espagnol, c'est-à-dire les langues qui n'exigent pas la présence d'un sujet lexical explicite, il n'existe aucun équivalent systématique du pronom *on*.

La variété des équivalents reflète aussi la variété des degrés de détermination de ce pronom. Comment est-il possible d'identifier le degré de détermination du référent du pronom *on* en s'appuyant sur ses équivalents en tchèque et en espagnol ? Afin de répondre à cette question principale, une étude du corpus sera effectuée. L'analyse sera basée sur le corpus InterCorp, plus précisément sur les sous-corpus des textes juridiques (Acquis) et des textes littéraires (Core, limité aux originaux français). Les informations essentielles pour la partie théorique seront puisées dans les grammaires et dictionnaires français, tchèques et espagnols ainsi que dans les journaux spécialisés en linguistique française.

Ce mémoire sera divisé en trois parties. La première partie analysera le pronom *on* du point de vue théorique, c'est-à-dire étymologique, syntaxique, stylistique et sémantique. De plus, les points de repère contrastifs vont être présentés afin de construire des bases solides pour l'analyse de corpus.

La deuxième partie introduira la méthodologie utilisée, surtout les corpus de travail et les conditions de la recherche, y compris l'expression régulière. En outre, afin de répondre à la question de recherche principale sur les l'indication du degré de détermination du pronom *on*, une série des questions de recherche partielles sera présentée de même que les hypothèses.

En ce qui concerne les hypothèses principales, il est probable que le degré de détermination du pronom *on* différera en fonction du type de texte, en supposant que



les textes juridiques auront recours plutôt à l'emploi indéterminé du pronom *on*. Etant donné que le tchèque et l'espagnol disposent de procédés formels d'impersonnalisation semblables, il est possible que le degré de détermination soit similaire dans les traductions aussi. En outre, le degré de détermination sera influencé par le verbe qui accompagne le pronom. La diversité lexicale des textes littéraires entraînera donc une diversité des degrés de détermination. Finalement, il est possible de supposer que c'est surtout le contexte qui peut indiquer le degré de détermination.

# 1 PARTIE THÉORIQUE

Commençons par citer une remarque prononcée par Păuna (2013) : « *Le pronom **on** est une forme grammaticale passe-partout ; il représente un moyen d'assurer l'économie dans le discours mais peut devenir un vrai piège pour les traducteurs* ». Cette phrase résume bien l'enjeu : écrire simplement un *on* au lieu de préciser par exemple *tous les hommes*, c'est non seulement une simplification et raccourcissement de l'énoncé, mais aussi sa condensation. En même temps, ce pronom a aussi le pouvoir de mettre l'énoncé dans un flou afin d'atteindre de divers objectifs : laisser le lecteur choisir son interprétation ou bien le tromper, créer du suspense, effacer l'auteur et beaucoup plus.

Pour ces raisons, ce chapitre vise à répondre aux questions concernant le fonctionnement du pronom *on*, ses référents et son usage dans de divers styles (plus précisément le style littéraire et juridique). De plus, cette analyse pourrait être intéressante aussi du point de vue contrastif car un des objectifs de cette étude est de comparer les équivalents du pronom *on* dans deux langues systématiquement divergentes. Aussi, étant donné qu'aucune grammaire n'établit une liste exhaustive des valeurs référentielles du pronom *on*, un tableau regroupant les différentes possibilités d'interprétation sera dressé.

## 1.1 Pronom *on*, défi à de nombreuses facettes

La problématique de la pluralité des perspectives dans les recherches sur le pronom *on* est très pertinemment introduite par la citation suivante de Blanche-Benveniste (2003, p. 43) : « *À examiner les emplois de **on** dans les usages les plus courants de la langue française contemporaine, il y a lieu de s'étonner de l'agilité des francophones, qui utilisent ce pronom dans des significations parfois opposées, en s'y embrouillant très rarement.* »

Ainsi, il est évident que le pronom *on* constitue un défi à plusieurs niveaux. D'abord, les points de repère stylistiques seront présentés afin de mieux encadrer le chapitre 3 en étant conscients des différences entre les textes littéraires et juridiques. Puis, l'étymologie ainsi que les usages historiques du pronom *on* seront analysés. Dans le cadre de ce sous-chapitre, une étude des règles pour l'insertion de l'article élide avant le pronom sera effectuée. Finalement, la problématique du degré de détermination sera mise en évidence en examinant les différentes approches à la classification grammaticale de ce pronom.

### 1.1.1 Pronom *on* du point de vue stylistique

Afin d'atteindre l'objectif d'examiner le degré de détermination du pronom *on* dans les traductions des textes divers, il est d'abord nécessaire d'observer les caractéristiques de chacun des deux styles en question : littéraire et juridique.

Parquet (1996, p. 38) caractérise la stylistique du discours législatif par son caractère abstrait, général et permanent. Vu qu'il s'agit d'un ordre ou bien d'un commandement, il a pour objectif d'exprimer une règle de droit universellement applicable. La généralité peut être exprimée par le sujet et c'est pour cela que, selon elle, le langage législatif recourt à l'usage de nombreux pronoms indéfinis.

Čechová, Krčmová et Minaříková (2008, p. 218–222) soulignent le caractère stéréotypique des textes juridiques qui créent les normes, en s'appuyant sur l'index des répétitions des mots plus élevé que dans les autres types de textes. En outre, selon elles, le vocabulaire a souvent recours à l'emploi des noms au détriment des verbes. La diversité lexicale des verbes reste limitée.

Le style des textes littéraires est difficilement définissable lors que chaque auteur choisit son propre mode de travail et il n'y a pas de règles strictes établies. Ainsi, il est possible que le pronom *on* apparaisse à plusieurs reprises car ne pas tout expliciter peut être désirable. L'enjeu de l'explicitation et de l'indétermination est évoqué aussi par Čechová, Krčmová et Minaříková (2008, p. 317–322). Elles insistent sur le fait qu'une œuvre littéraire exige une activité interprétative soignée de la part du lecteur. En ce qui concerne le vocabulaire, l'index des répétitions des mots est moins élevé que dans les autres types de textes, ce qui confirme que la richesse lexicale est supérieure à celle des textes juridiques.

Cependant, lors de l'étude de la prose comprise dans le corpus InterCorp, il faut bien distinguer la zone du narrateur de celle des personnages. En outre, le narrateur peut être omniscient, l'analyse des textes littéraires exige donc un travail plus rigoureux au niveau du contexte.

Ces observations seront cruciales pour le chapitre 3 – il est possible que les textes littéraires offrent une vaste palette des verbes qui accompagnent le pronom *on*, les textes juridiques auront plutôt tendance à réduire la variabilité de ces verbes.

### 1.1.2 Usages historiques du pronom *on* et l'insertion de « l' »

La première occurrence du pronom *on* (sous la forme de *om*) attestée à l'écrit a apparu dans les Serments de Strasbourg : « *si cum om per dreit son fradra salvar dif* ». Au fur et à mesure, il a été employé dans la Chanson de Roland ou chez Racine et Molière. Selon le *Dictionnaire du français classique* (1971, p. 389), le pronom *on* s'employait souvent au XVII<sup>e</sup> siècle pour la première personne afin d'éviter le « *moi haïssable* ». Il s'agissait donc d'une marque de modestie. Néanmoins, Bouhours (1671, p. 122) précise les conditions de son emploi en disant que : « **On** se dit à toute heure dans un sens nouveau. Car pour dire : « *Je vous en serai obligé, je ferai mon devoir, n'oubliez pas au moins ce que je fais pour vous* », nous disons en parlant et en écrivant familièrement aux personnes qui nous sont égales ou inférieures : « **on** vous en sera obligé, **on** fera son devoir, n'oubliez pas au moins ce qu'**on** fait pour vous ». Ce ne serait pas être juste dans le langage, que d'user de cette expression à l'égard des personnes qui sont au-dessus de nous. » Il en résulte que dans l'histoire, de même que maintenant, le référent du pronom *on* demeurait vague.

La *Grammaire Larousse du français contemporain* (1997, p. 230) dit que grâce au caractère nominal du pronom, son ancien article peut être gardé « *selon de subtiles règles d'harmonie précisées au XVII<sup>e</sup> siècle (à l'initiale, après voyelle, dans si l'on, où l'on, que l'on...)* ». Hendrich, Radina et Tláškal (2001, p. 324) soulignent son caractère littéraire. De plus, cette forme apparaît surtout après les conjonctions *si*, *et*, *que* et les pronoms relatifs *que* et *où*.

Le *Dictionnaire des difficultés du français* (1999, p. 503) ajoute une explication plus précise de ce fait : le « l' » se place parfois devant *on* par souci d'harmonie, surtout pour éviter la prononciation [kɔ̃]. Cependant, cet emploi n'est pas obligatoire. Les règles d'euphonie ont été établies par Vaugelas afin d'éviter le hiatus après les conjonctions déjà mentionnées. A l'époque contemporaine, Narjoux (2010, p. 36–45) souligne que l'emploi de l'article élide dénote un niveau de langage soutenu. Il est donc plus probable que lors de notre recherche dans le corpus, l'article élide n'apparaîtra pas très souvent. Pourtant, vu qu'il ne s'agit pas d'un aspect négligeable, il faut en tenir compte lors de l'analyse de corpus et surtout lors de la recherche (afin de ne pas effacer une partie des résultats).

Pour résumer cette remarque sur les usages historiques et l'insertion du pronom *on*, citons Grevisse et Goose (2008, p. 648). Ils insistent sur le fait qu'autrefois, le pronom *on*

permettait que le référent du pronom *on* varie dans une même phrase (il pouvait donc dénoter plusieurs référents). Cette fonction n'a toutefois pas persisté et le pronom *on* en français contemporain ne désigne qu'une seule entité humaine à la fois. Ce qui relie les usages contemporain et historique, c'est en tout cas la diversité de référents, dont fait preuve aussi l'impossibilité de se mettre d'accord sur la catégorisation grammaticale du pronom *on*.

### 1.1.3 Appartenance du pronom *on* à la classe de mots

Pour démontrer que ce n'est pas seulement un souci au niveau de la traduction, mais aussi au niveau de la catégorisation, trois perspectives des parties du discours vont être présentées. Ce débat n'est pas inutile puisque le pronom *on* peut acquérir de nombreuses fonctions qui peuvent provoquer des changements radicaux du sens de la phrase (voir le chapitre 1.2 traitant sur ses valeurs référentielles).

Ainsi, pour clarifier l'enjeu de la détermination du référent du pronom *on*, il faut chercher les arguments pour que le pronom *on* puisse être classifié parmi les pronoms indéfinis, personnels et enfin personnels indéfinis.

D'abord, la *Grammaire Larousse du français contemporain* (1977, p. 3772) classe le pronom *on* parmi les pronoms personnels en soulignant qu'il s'agit d'un cas particulier. Dubois (1965, p. 111–112) développe cette idée en réclamant que « *les référents personnels [...] supposent la négation du système, c'est-à-dire un segment qui soit exploité comme la suppression de la personne [...]; le pronom on, qui remplit ces fonctions, est donc un substitut de la même classe que je, tu, il* ». Pour lui, la capacité combinatoire inférieure constitue la seule différence entre *on* et les autres pronoms personnels.

Boutet (1986, p. 31) aussi refuse de le classer parmi les indéfinis en problématisant l'enjeu. Elle donne deux exemples :

(1) « *Hommes et femmes **on** est payé pareil.* »

(2) « *Ah ben on dit toujours si **on** avait pu.* »

Ces exemples illustrent bien que l'usage anaphorique (exemple 1) dispose d'une référence nettement définie, alors que la référence du deuxième exemple généralisant est indéfinie. Bien sûr, ces explications ne suffisent pas pour observer la nature de ce pronom.

Atlani (1984, p. 14) résume bien les incohérences des opinions : « *Les tentatives qui réduisent **on** à un indéfini ou à un pronom personnel bloquent l'analyse.* » Or, le sujet doit

être problématisé en mélangeant les deux catégories, les pronoms personnels et les pronoms indéfinis. Citons par exemple le *Dictionnaire des difficultés du français* (1999, p. 504–505) : « ce pronom « indéfini » est en réalité un pronom personnel se substituant en fonction du sujet à n'importe lequel des mots *je, tu, il, elle, nous, vous, ils, elles* ».

En bref, ces thèses dégagent l'ambiguïté du pronom *on* non seulement au niveau syntaxique, mais aussi morphologique. Ainsi, cette ambiguïté morphologique reflète aussi l'ambiguïté sémantique et référentielle qui sera étudiée dans le chapitre suivant.

## **1.2 Détermination du référent du *on***

Etant donné que l'enjeu principal de ce mémoire, c'est d'examiner les différents degrés de détermination du référent du pronom *on* et les facteurs qui les influencent, il faut d'abord s'interroger sur les différentes valeurs référentielles du pronom *on*. Puis, l'accord de l'adjectif ou le participe passé avec le verbe sera examiné car il peut aider à déterminer la quantité et la qualité du référent (c'est-à-dire le nombre et le genre du référent).

### **1.2.1 Valeurs référentielles du pronom *on***

Pour interpréter le pronom *on*, il faut recourir au contexte non seulement immédiat, mais aussi plus large. Ceci peut constituer une des sources d'ambiguïté référentielle de ce pronom. Vu que la variété des valeurs référentielles du pronom *on* joue un rôle considérable lors de la traduction en tchèque et en espagnol, comprendre le fonctionnement discursif de *on* est essentiel pour ce travail. C'est pour cela que dans ce chapitre, les différents référents du pronom *on* vont être dégagés et concrétisés à l'aide des recherches faites par de nombreux linguistes.

Le premier élément important est observé par la *Grammaire du français classique* (2002, p. 47) : le référent du pronom *on* est toujours animé. Puis, une vaste palette des référents s'ouvre. Landragin et Tanguy (2014) citent plusieurs aspects qui influencent la détermination du référent. Il s'agit d'abord de la morphologie ; la présence des pronoms personnels tels que *nous* peut indiquer qu'il s'agit d'un usage déterminé. Puis, ce sont des facteurs sémantiques. Finalement, l'usage joue un rôle important aussi (par exemple l'usage

dans les proverbes). Bref, le contexte intra-phrastique de même que le contexte extra-phrastique demeurent essentiels lors de l'interprétation du référent du pronom *on*.<sup>1</sup>

Voici un tableau qui résume les différentes valeurs référentielles du pronom *on*.

Degré	Exemple	Commentaires	Source
DÉTERMINATION	(1) <i>Ces notes furent écrites en 1910. <b>On</b> était fort loin de penser qu'<b>on</b> les donnerait enfin au public.</i>	Marque de modestie, remplacement de « je » de l'auteur	<i>Dictionnaire des difficultés du français (1999)</i>
	(2) <i>Dans cet article, <b>on</b> se propose de démontrer [...].</i>	Remplacement de « je » de l'auteur dans les articles scientifiques	<i>Grand Larousse de la langue française (1971)</i>
	(3) <i>Elle est offensée, il rit : « <b>On</b> est fâchée ? Elle est fâchée » dit-il en s'adressant à moi.</i>	Ironie, familiarité	<i>Grammaire du français contemporain (1997)</i>
	(4) <i>A-t-<b>on</b> été sage, aujourd'hui ?</i>	Affection, sympathie, enjouement ou mépris, emploi hypocoristique	<i>Grammaire critique du français (2007)</i>
	(5) <i><b>On</b> tient essentiellement, ma mère – <b>on</b>, c'est mon client.</i>	Dénotation d'une tierce personne	<i>Dictionnaire des difficultés du français (1999)</i>
	(6) <i>Si <b>on</b> était malins, <b>on</b> le garderait pour nous, le studio de Patrick.</i>	Familier et populaire, mais « nous » préférable <sup>2</sup>	<i>La grammaire des fautes (1929)</i>
VÉRITÉ GÉNÉRALE	(7) <i>Quand <b>on</b> veut, on peut.</i>	Proverbes, être humain non précisé	<i>Dictionnaire du français contemporain (1971)</i>
	(8) <i>Autrefois, <b>on</b> vivait mieux.</i>	Détachement temporel et spatial	<i>Grammaire critique du français (2007)</i>
INDÉTERMINATION	(9) <i>Gardes, qu'<b>on</b> m'obéisse !</i>	Dénotation d'un ensemble de personne imprécis	<i>Grammaire critique du français (2007)</i>
	(10) <i><b>On</b> a barré la route.</i>	Correspond au passif sans complément d'agent	<i>Dictionnaire du français contemporain (1971)</i>
	(11) <i><b>On</b> vous demande à la loge.</i>	Personne dont l'identité n'est pas indiquée (indifférence, ignorance)	<i>Dictionnaire du français contemporain (1971)</i>

**Tableau 1 : Valeurs référentielles du pronom *on* selon le degré de détermination**

<sup>1</sup> Gjesdal (2009) observe trois niveaux de contexte et leur influence sur la référence du pronom *on*. Cette étude découvre que ce ne sont pas seulement les marqueurs temporels, mais aussi spatiaux qui déterminent le référent. De plus, le niveau sémantique est très important. C'est surtout ce niveau qui sera étudié dans le chapitre 3.1.

<sup>2</sup> Wilmet (2007, p. 213) trouve que l'emploi de *on* à la place de *nous* est même un vulgarisme, par exemple : « **On** a d'abord espéré ça, nous aussi. » (Jean Anouilh, *Antigone*)

Etant donné qu'aucune grammaire n'offre une liste exhaustive des valeurs référentielles du pronom *on*, les données contenues dans ce tableau sont puisées dans de nombreux ouvrages. Il est possible d'en dégager une tendance générale : les valeurs peuvent être regroupées en trois degrés de détermination.

Ainsi, le pronom *on* peut désigner des entités plus ou moins bien définies, dont font preuve les exemples 1–6.<sup>3</sup> Le référent dans les exemples 1 et 2 est semblable : il s'agit de « *je* » remplacé par le « *on* » pour de différentes raisons (modestie, usage). Les exemples 3 et 4 introduisent un *on* hypocoristique qui décrit une personne précise avec l'ironie ou la familiarité. Le référent dans les exemples 5 et 6 est précisé dans le contexte immédiat, c'est donc pour cela que lors de l'étude du corpus, le contexte joue un rôle très important.

Le deuxième degré de détermination est représenté par les vérités générales. Le « *on* » semble atemporel, comme dans l'exemple 7, il est souvent utilisé dans les proverbes et les dits populaires. L'exemple 8 est encore moins déterminé, ce qui témoigne du fait qu'un classement indiscutable des exemples demeure impossible.

Finalement, les exemples du troisième degré, c'est-à-dire de l'indétermination, ne prêtent pas attention au référent. Néanmoins, comme dans l'exemple 8, l'exemple 7 ne semble pas totalement indéterminé, c'est surtout le contexte qui peut aider à décider si le « *on* » est totalement indéterminé, ou bien s'il s'agit d'une dénotation de *vous* ou *ils*. Ceci donc prouve que les limites des degrés de détermination ne sont pas nettes. En revanche, les exemples 10 et 11 effacent totalement la détermination du référent.

Ces observations seront très utiles surtout dans les chapitres 3.2.2, 3.3.2 et 3.4 où les échelles de la détermination du pronom *on* seront construites et les mêmes degrés seront observés et comparés en tchèque et en espagnol dans les textes littéraires et juridiques.

### 1.2.2 Accord de l'adjectif ou du participe passé

C'est aussi l'accord de l'adjectif ou du participe passé avec le verbe qui peut jouer un rôle assez important lors de la détermination du référent du pronom *on* car en présence de l'accord, la quantité ou le nombre (et d'ailleurs la qualité ou le genre) des référents sont

---

<sup>3</sup> Bien sûr, bien que le nombre d'exemples de détermination excède le nombre d'exemples des autres degrés, cela ne signifie pas qu'il soit le plus représenté dans le corpus. Or, les référents des exemples qui rentrent dans la catégorie de la détermination sont plus variés que les référents des autres catégories.



évoqués sans que le traducteur doive faire face à l'interprétation du contexte extralinguistique. Comme le souligne le *Dictionnaire des difficultés du français* (1999, p. 503), l'adjectif et le participe passé peuvent prendre les marques de nombre et de genre (et donc ne pas rester invariables, mais tout de même ne sont pas obligés de varier non plus).

En français classique (plus précisément au XVII<sup>e</sup> siècle), le pronom *on* exigeait un accord au masculin singulier (Fournier, 2002, p. 47) : « **L'on** est étonné, avec tout son esprit, de se trouver la dupe de plus sots que soi » (La Bruyère, *Les Caractères*, 1688), parfois, selon le contexte au féminin : « Dans le temps où l'on est belle » (Molière, *La Princesse d'Élide*, 1664) ou au pluriel : « Et l'on ne devient guère si riches à être honnêtes gens » (Molière, *Le Bourgeois Gentilhomme*, 1670).

Hendrich, Radina et Tláškal (2001, p. 324) disent que lorsque le pronom désigne un ensemble des femmes ou, en général, des personnes, l'adjectif s'accorde *ad sensum* : « Quand on est jeune et belle... ». Ainsi, il est possible de réduire le champ des références possibles au singulier ou au pluriel, il s'agit donc d'un procédé qui aide à la désambiguïsation.

### 1.3 Répertoire des équivalents du pronom *on* en tchèque et en espagnol

Le français, à la différence des autres langues romanes, n'a pas la capacité d'omettre le sujet (*null-subject language*).<sup>4</sup> De l'autre côté, l'expression explicite du sujet n'est pas obligatoire en tchèque ou en espagnol. Ces deux langues disposent de traits grammaticaux et phonétiques assez univoques pour distinguer les différentes personnes verbales sans mentionner explicitement le sujet.

De réussies études ont été élaborées sur cette problématique, par exemple Dušková (1999), Čermák (2010) ou Tláškal (2005). Pourtant, les linguistes espagnols ne s'y intéressent pas beaucoup, dont témoigne le manque d'articles dans des journaux scientifiques ou des monographies.

---

<sup>4</sup> Pourtant, Grevisse et Goose (2008, p. 667) indiquent que le français n'a pas toujours été une langue nécessitant l'expression explicite du sujet. *Le Bon Usage* présente l'effacement de la diversité des terminaisons verbales comme la source de la nécessité d'insérer le sujet. Au Moyen Âge, le pronom sujet était assez souvent omis car les terminaisons verbales étaient encore sonores et donc ne nécessitaient pas une précision grammaticale de la personne.

Le problème auquel il faut faire face lors du travail avec des grammaires franco-tchèques et franco-espagnoles, c'est surtout leur tendance à introduire une liste des équivalents du pronom *on* sans les classer ou regrouper. Soit les auteurs renoncent complètement à prendre en considération le contexte et citent des exemples peu nombreux desquels le lecteur ne peut déduire aucun système, comme Fernández Ballón et Monnerie-Goarin (1994) ; soit ils essaient d'introduire plusieurs possibilités de traduction, malheureusement sans ambitionner de les systématiser, comme Radina (1977).

C'est pour cela qu'un des objectifs de ce mémoire est aussi de proposer une catégorisation complète des valeurs référentielles car une telle catégorisation n'est offerte par aucun dictionnaire ou grammaire. Puis, les valeurs référentielles seront examinées à partir des traductions en tchèque et en espagnol vu que ces deux langues ne disposent pas d'un équivalent direct du pronom *on*. Finalement, les équivalents seront regroupés en trois degrés de détermination (pour plus de détails, voir les chapitres 3.2.2 et 3.3.2).

### 1.3.1 Équivalents du pronom *on* en espagnol

Dans les manuels scolaires, le sujet de la traduction du pronom *on* reste très peu abordé. Considérons quelques exemples. Fernández Ballón et Monnerie-Goarin (1994, p. 75) se contentent, tout en classant le pronom *on* parmi les pronoms indéfinis, de l'explication peu complète : « *[Le pronom] on fait référence à un ensemble de personnes indéterminé.* » Pour soutenir cette thèse, ils prennent l'exemple de la phrase « *on a toujours besoin d'un plus petit que soi* » dont le sujet interprété, selon les auteurs, est la troisième personne du pluriel. Ensuite, ils mentionnent l'usage typique de la langue parlée qui indique la première personne du pluriel (« *on a peur* ») ainsi que les usages affectifs (« *Alors, on travaille !* ») qui peuvent être interprétés comme la deuxième personne du singulier ou pluriel. Enfin, les auteurs ajoutent pertinemment que *on* est toujours suivi d'un verbe à la troisième personne du singulier.

Alloa et De Torres (2005, p. 104–105) se mettent d'accord sur la catégorisation du pronom *on* parmi les pronoms indéfinis. Néanmoins, ils insistent sur de nombreuses spécificités du pronom *on* qui le distinguent des autres indéfinis : selon eux, il ne dénote ni la totalité, ni l'extraction aléatoire, ni la quantité nulle. Ainsi, ils interprètent

cette variabilité des référents comme une marque de ce qu'ils appellent « *aréférentialité* » (c'est-à-dire l'absence du référent).<sup>5</sup>

Zavadil et Čermák (2010, p. 437–438) distinguent 2 formes du sujet générique : (1) la troisième personne du pluriel (si le locuteur est exclu) ; (2) la deuxième personne du singulier (si le locuteur est inclus). De plus, l'impersonnalité peut être exprimée par le passif réfléchi (le pronom réfléchi *se* fait partie du prédicat dont le nom est l'objet) : « *Se desprecia a la mujer* » ou par une phrase passive où le sujet est identique au patient : « *La gramática se estudia mucho* » (Zavadil et Čermák, 2010, p. 441–442). Enfin, pour mieux comprendre les possibilités de traduction en espagnol, il faut puiser dans des articles spécialisés qui, pourtant, restent assez peu nombreux.

Yao (2016, p. 124–127) met en avant l'usage de *uno* comme un pronom indéfini avec un référent indéfini ou générique. Dans ce cas, elle considère que le pronom a perdu sa valeur déictique personnelle et peut inclure ou exclure le locuteur : « *A mi edad, uno no puede valerse por sí mismo* ». A la différence des autres, elle suggère aussi l'emploi des pronoms personnels *yo* et *tú* substituant *uno* afin de ne pas indiquer un référent concret. Parmi les autres possibilités d'équivalents du pronom *on*, il est possible de classer la troisième personne du pluriel et *se* + verbe. En ce qui concerne le langage familier, elle propose la traduction par la première personne du pluriel. De l'autre côté, Akrobou (2011, p. 45) n'admet que deux possibilités de traduction du pronom *on* en espagnol : par la troisième personne du pluriel et par le « *se* » réfléchi.

Atlani (1984, p. 16) cite des exemples de l'usage du pronom *on* tout en soulignant difficultés que nous pouvons rencontrer lors de la traduction en espagnol. D'abord, elle accentue que le singulier en espagnol, sans aucune précision, n'a pas de valeur indéfinie que le pronom *on* désigne. Puis, elle insiste sur la nécessité d'examiner le contexte de l'énoncé pour déterminer si le pronom fait référence à la première ou troisième personne du pluriel (c'est-à-dire *nous* inclusif ou *ils* exclusif). Cette observation est essentielle pour le chapitre 3 dans lequel les données seront puisées dans le corpus InterCorp afin de dégager le fonctionnement de ce pronom.

---

<sup>5</sup> L'aréférentialité est un terme assez problématique. Les chapitres 3.2, 3.3 et 3.4 vont essayer de contredire cette notion car bien que le référent soit indéterminé, il n'est pas totalement supprimé. Ainsi, un meilleur terme que « *l'absence du référent* » devrait être employé, il s'agit plutôt de « *l'absence du référent explicite* ». En outre, pas toutes les occurrences du pronom *on* correspondent à cette thèse du référent indéterminé.

Pour résumer cette partie, à partir des avis des philologues et traducteurs de la langue espagnole, un tableau qui servira de base pour la recherche qualitative présentée dans les chapitres 3.2 et 3.3 peut être dressé.

Type		Exemple
Type lexical ( <i>alguien/uno/nadie</i> etc.)		<i>En España, <b>la gente</b> se acuesta tarde.</i>
Type morphosyntaxique	1 <sup>ère</sup> personne du pluriel	<i>Nos <b>vamos</b>.</i>
	2 <sup>e</sup> personne du singulier/pluriel	<i>Si <b>estás</b> triste, conviene desahogarse con los amigos.</i>
	3 <sup>e</sup> personne du pluriel	<i><b>Vienen</b> a recoger la ropa usada.</i>
Type passif sans complément d'agent (passif promotionnel)		<i>No me gusta que mi libro <b>sea tomado</b> a la ligera.</i>
Type passif avec <i>se</i> réfléchi (passif réfléchi)		<i><b>Se</b> necesita a ese ayudante.</i>
Type nominalisation	Nominalisation subordonnée sujet –	<i>Muere <b>a tiros</b> de carabina.</i>
	Nominalisation subordonnée temporelle –	<i><b>Su rubor</b> significaba una respuesta afirmativa.</i>

**Tableau 2 :** *Variété des équivalents du pronom on en espagnol*

Cependant, ce tableau basé sur les constatations de Tláskal et d'autres linguistes ne correspond pas totalement aux besoins du présent mémoire. Alors qu'ils s'intéressent plutôt au classement des types plutôt grammatical, ce mémoire veut travailler avec des types d'équivalents regroupés en degrés de détermination.

De ce point de vue, l'analyse la plus soignée se trouve dans le dictionnaire franco-espagnol Larousse qui divise les significations de *on* en 4 parties :

(1) *nous* (familier) qui se traduit par la première personne du pluriel

« **on** *s'en va* » = « *nos vamos* »

(2) quelqu'un qui se traduit par la troisième personne du singulier ou pluriel, facultativement avec la traduction littérale *alguien*

« **on** *t'a téléphoné ce matin* » = « *te ha llamado **alguien** esta mañana / esta mañana te **han** llamado* »

(3) les gens, l'espèce humaine, groupe de personnes qui se traduit par la gente ou troisième personne du pluriel

« *en Espagne, **on** se couche tard* » = « *en España, **la gente** se acuesta tarde* »  
« ***on** raconte / **on** dit que* » = « ***dicen** que* »<sup>6</sup>

(4) sujet indéterminé qui se traduit par *se*

« ***on** n'a pas le droit de fumer ici* » = « *aquí no **se** puede fumar* »  
« ***on** ne sait jamais* » = « *nunca **se** sabe* »

Voici une première ébauche de l'échelle de la détermination qui sera présentée d'une manière plus détaillée dans les chapitres 3.2.2, 3.3.2 et 3.4. Le premier exemple dénote un référent plus ou moins précis alors que les exemples 2 et 3 sont plus généraux. Le dernier exemple est l'indétermination totale du référent. Cette observation servira d'un point de départ pour l'analyse dans la partie empirique.

### 1.3.2 Equivalents du pronom *on* en tchèque

De l'autre côté, la langue tchèque dispose de moins de traits systémiques communs avec le français. Pourtant, de même que la grammaire espagnole, la grammaire tchèque ne présente aucun équivalent direct du pronom *on*. Ainsi, le même défi surgit : celui de le traduire de la façon la plus pertinente possible.

Voici donc, en premier lieu, le principe en tchèque. Pour des raisons assez compréhensibles, c'est-à-dire la diversité énorme des référents du pronom *on* et la richesse de langue ainsi que la diversité des registres littéraires, aucune grammaire ou dictionnaire n'établit une liste définitive de ses traductions.

*Le dictionnaire franco-tchèque* de Neumann, Hořejší, Vlasák et Kolář (1992, p. 188) introduit une liste des possibilités de traduction du pronom *on* : *člověk, někdo* et *jeden, my, lidé*, des formes verbales, telles que les formes passives réfléchies contenant *se* ou le verbe *být* suivi d'un participe, ou bien les personnes grammaticales. De plus, il suggère que

---

<sup>6</sup> Remarquons pourtant que le corpus CORPES de la Real Academia Española (2018) documente la préférence d'utiliser le verbe *decir* à la troisième personne du pluriel. La variante *se dice que* est pourtant toujours visible, bien qu'elle soit quatre fois moins représentée.

le pronom *on* peut remplacer, dans le français parlé, le pronom personnel de la première personne ou de la deuxième personne.

Radina (1977, p. 168–169), sous forme d'exercices guidés, introduit une quinzaine d'exemples. Il ne les regroupe pas, il est toutefois possible de les classer en 5 catégories : (1) passivation en utilisant le pronom réfléchi *se* ; (2) inclusion du locuteur en utilisant la première personne du pluriel ; (3) exclusion du locuteur en utilisant la troisième personne du pluriel ; (4) inclusion de l'interlocuteur en utilisant la deuxième personne du singulier ou pluriel ; (5) généralisation à l'aide des noms tels que *člověk* (*l'homme*), *někdo* (*quelqu'un*) ou *jeden* (*un*).<sup>7</sup>

Hendrich, Radina et Tláskal (2001, p. 323–324) consacrent plus d'une page au pronom *on*. Selon eux, il s'agit d'un pronom indéfini personnel clitique qui désigne uniquement les êtres humains. Parmi les exemples de traduction en tchèque, ils citent *člověk* (*l'homme*), *někdo* (*quelqu'un*), *lidé* (*les gens*), puis les constructions passives avec le pronom réfléchi *se* et la voix passive *být* + participe (*être* + participe) et d'autres personnes verbales du singulier tout comme du pluriel.

Pour ce qui est de l'expression de l'impersonnalité en tchèque en général, Grepl et Karlík (1998, p. 133–140) introduisent le terme « *désagentisation* » pour le phénomène qui est assez proche de l'emploi du pronom *on* en français. L'agent est repoussé de la position du sujet et c'est l'actant substantiel qui prend sa place, par exemple : « *Obecní úřad přidělil mladým manželům malou garsoniálu.* → *Mladým manželům byla přidělena malá garsoniéra.* » Ils constatent que les procédés principaux de ce procès sont grammaticaux, plus précisément : (1) passif périphrastique ; (2) *se* réfléchi + verbe ; (3) troisième personne du pluriel ; (4) d'autres personnes grammaticales. Puis, ils diversifient aussi la quatrième catégorie : (4.1) deuxième personne du singulier dénotant possibilité, capacité ; (4.2) première personne du pluriel dénotant nécessité ; (4.3) première personne du singulier dénotant reproche, peu fréquente ; (4.4) troisième personne du singulier. L'emploi des procédés lexicaux, selon eux, reste marginal – par exemple « *Maminka dala Petrovi na Vánoce kolo.* → *Petr dostal od maminky na Vánoce kolo.* » Bien qu'il s'agisse d'une

---

<sup>7</sup> Le travail d'Obradović offre une liste assez riche des moyens comment traduire *on* en une langue slave, plus précisément en serbe. Elle mentionne les équivalents *неко* (*quelqu'un*), *човек* (*l'homme*), *свет зна* (*le monde sait*), *људи знају* (*les gens savent*), *један* (*un*) ou *се* (*se*). De plus, Obradović en citant Drašković fait une remarque intéressante : bien qu'en français familier, le pronom *on* puisse désigner les trois personnes, le serbe ne permet pas la traduction en deuxième personne.

œuvre concentrée seulement sur la langue tchèque, les tendances de traduction du pronom *on* en peuvent être tirées. Ainsi, ce sont aussi Grepl et Karlík qui permettront l'établissement de la classification des équivalents du pronom *on* introduite dans le chapitre 3.2. et 3.3.

Bien sûr, il ne faut pas oublier les articles spécialisés qui nous offrent une base assez solide pour la deuxième partie de ce mémoire, c'est-à-dire l'analyse du corpus. Le travail de Tláskal (2005, p. 380–383) repose sur l'analyse du corpus plurilingue de l'œuvre *Le petit prince* d'Antoine de Saint-Exupéry. Il a conscience de la nature différente du tchèque et de l'espagnol d'un côté (les langues à sujet nul) et du français de l'autre côté (où la présence d'un sujet explicite est obligatoire). Tláskal en tire la conclusion qu'il existe six types d'équivalents du pronom *on* : lexical, morphosyntaxique, passif avec *se* réfléchi, structure à sujet nul, nominalisation, conversion syntaxique (voir le tableau 3).

Beneš (1965, p. 172) toutefois refuse l'équivalence entre *on* et *člověk* à valeur d'un sujet indéterminé (que Tláskal classe parmi les traductions lexicales) car *člověk* n'est pas « *un instrument si commode et fréquent comme l'est **on** en français* ». De l'autre côté, Čermák (2010) élabore une étude qui observe les correspondances entre *on* et *člověk* en s'appuyant sur le corpus InterCorp (dans la combinaison français-tchèque avec *Double vie* de Pierre Assoulin). Il observe si le tchèque préfère la traduction *člověk* du pronom *on* et conclut que cette solution ne constitue qu'un quart des équivalents. Selon Čermák, la majorité des équivalents sont des équivalents grammaticaux (notamment le *se* réfléchi et la troisième personne du pluriel) et des cas à part. Ce sont surtout ces deux études de Tláskal et Čermák et leurs classifications qui serviront de base pour le tableau suivant.

Type		Exemple
Type lexical		<i>Je velice užitečné, když <b>člověk</b> v noci zabloudí.</i>
Type morphosyntaxique	1 <sup>ère</sup> personne du pluriel	<i>Nikdy <b>nejsme spokojení</b> tam, kde <b>jsme</b>.</i>
	2 <sup>e</sup> personne du singulier/pluriel	<i>Nikdy <b>nevíš</b>.</i>
	3 <sup>e</sup> personne du pluriel	<i>Abych mohl zdravít, když mě s jásotem <b>vítají</b>.</i>
Type structure à sujet nul		<i>Na Zemi <b>je vidět</b> všelicos.</i>
Type passif avec <i>se</i> réfléchi		<i>Vypravování badatelů <b>se</b> zaznamenávají nejprve tužkou.</i>
Type nominalisation	Nominalisation – subordonnée temporelle	<i>Inkoustem se zapíše, <b>teprve</b> až badatel podá důkazy.</i>
Type conversion syntaxique	Conversion sujet x objet	<i>Jak <b>nám</b> mohou patřit hvězdy?</i>

**Tableau 3 :** Variété des équivalents du pronom *on* en tchèque

Ce tableau représente surtout le point de vue de Tláskal et Čermák. Il faut toutefois prendre en considération qu'une approche scalaire servirait mieux à dégager le fonctionnement du pronom *on* plutôt qu'une énumération des traductions. De plus, ces types ne correspondent pas totalement aux besoins de ce mémoire vu que l'un des objectifs est de classer les référents selon le degré de détermination. Ainsi, il faudra retravailler ce tableau, compléter et regrouper le répertoire des équivalents.



## 2 PARTIE MÉTHODOLOGIQUE

La deuxième partie de ce mémoire a pour but de présenter la méthodologie appliquée lors de l'étude du pronom *on*. D'abord, il faut souligner que la présente étude est à la fois quantitative et qualitative car les corpus utilisés permettent d'effectuer des recherches des deux types.

C'est pour cela que, dans un premier temps, ces corpus vont être présentés. Après, les conditions de recherche ainsi que l'expression régulière utilisée et la méthode de traitement des données vont être définies. Ensuite, des limites de ce travail et les écueils rencontrés vont être dévoilés. Finalement, une série des questions de recherche et les hypothèses clé pour ce mémoire vont être introduites.

### 2.1 Présentation des corpus utilisés

La base des données sur laquelle est basée la présente étude, c'est le corpus InterCorp. Il s'agit d'un corpus parallèle, il permet donc d'effectuer des recherches contrastives. Cette analyse travaille avec la version 10 du corpus InterCorp qui a été lancée le premier décembre 2017.

Les données pour l'analyse des textes littéraires sont tirées du sous-corpus Core avec trois versions alignées : la française, la tchèque et l'espagnole. En outre, seuls les originaux français sont pris en compte et la recherche est aussi limitée aux textes de *fiction*. Ainsi, ce sous-corpus contient 7 161 043 tokens dans 9 œuvres (*Voyage au bout de la nuit* de Louis-Ferdinand Céline, *La Peste* et *L'Etranger* d'Albert Camus, *99 francs* de Frédéric Beigbeder, *Les Particules élémentaires* de Michel Houellebecq, *Les Mots* de Jean-Paul Sartre, *J'irai cracher sur vos tombes* de Boris Vian, *Le Petit prince* et *Lettre à un otage* d'Antoine de Saint-Exupéry).

Le sous-corpus Acquis sert de base pour la recherche dans les textes juridiques. Il s'agit d'un corpus regroupant les textes juridiques de l'Union européenne qui est majoritairement composé de textes des directives, règlements et d'autres actes législatifs. Contrairement au sous-corpus Core, dans Acquis, il n'est toutefois pas possible de distinguer la langue source. Il contient 31 274 516 tokens.

## 2.2 Conditions de la recherche

Le pronom *on* est un pronom clitique qui ne peut être séparé du verbe que par les pronoms suivants : *ne (n')*, *en*, *y*, *le (l')*, *la*, *les*, *lui*, *leur*, *me (m')*, *te (t')*, *se (s')*, *nous*, *vous*. Néanmoins, il ne faut pas oublier qu'il est possible de placer plusieurs mots entre les deux mots des formes verbales composées, par exemple les adverbes ou la négation :

- i. *Il y a longtemps qu'on y a renoncé.*
- ii. *Ainsi qu'on l'a déjà expliqué, les prévisions de rendement à long terme, qui peuvent s'écarter du rendement effectivement enregistré au cours d'années écoulées, constituent un facteur déterminant.*
- iii. *Ainsi, on n'a guère avancé sur le chemin de la réduction des émissions de gaz à effet de serre, en dépit des progrès accomplis à la fin des années 90.*

C'est pour cela qu'il faut tenir compte d'autres conditions limitant la recherche. Prenons l'exemple du passé composé : ce temps est composé de deux formes verbales, c'est-à-dire le verbe auxiliaire et le participe passé. Le moteur de recherche se contente d'indiquer le verbe *avoir* et *être* en tant que lemme recherché, même dans le cas où s'il s'agit d'un verbe auxiliaire. C'est pour cela qu'il faudra procéder à un tri manuel des résultats de recherche contenant ces verbes afin de pouvoir décider s'ils remplissent la fonction de verbe auxiliaire ou verbe à sens plein.

Il est pertinent d'utiliser l'expression régulière suivante :

```
[word="[Oo]n" & tag="PRO.*"] [tag="VER.*"] |  
[word="[Oo]n" & tag="PRO.*"] [] [tag="VER.*"] |  
[word="[Oo]n" & tag="PRO.*"] [] {2} [tag="VER.*"] |  
[word="[Oo]n" & tag="PRO.*"] [] {3} [tag="VER.*"]
```

Elle permet de rechercher le mot *on* (word) en tant que pronom (tag). Afin d'éviter de perdre les résultats qui seraient grammaticalement mal étiquetés, il faut prendre en considération tous les types de pronoms. C'est pour cela que l'expression régulière contient **tag="PRO.\*"** où le point signifie un caractère quelconque et l'astérisque dénote un nombre quelconque de répétitions. De plus, ce choix n'est pas aléatoire et ne compte pas qu'avec l'erreur d'étiquetage. Le chapitre 1.1.3 a montré les soucis de catégorisation

grammaticale : le pronom *on* est-il plutôt indéfini ou personnel ? L'expression régulière citée ci-dessus permet que l'interface trouve les deux interprétations, les résultats ne seront donc pas influencés par l'incohérence de l'étiquetage grammatical.

La deuxième partie de l'expression régulière introduit une, deux ou trois positions insérées entre le pronom et le verbe. L'expression ne compte qu'avec trois positions insérées puisque le verbe conjugué doit être situé le plus près possible du pronom *on*. Le plus près possible ne veut pas indispensablement dire en voisinage immédiat. Les différentes possibilités de mots insérés ont déjà été mentionnées, pour les résumer, il suffit de dire qu'entre *on* et le verbe conjugué peuvent apparaître des pronoms personnels, des pronoms adverbiaux et la négation.

Bien sûr, une expression régulière plus économique pourrait être appliquée. Ce n'est toutefois que celle-ci qui élimine les possibles erreurs. Par exemple, l'expression régulière `[word=" [Oo]n" & tag="PRO.*" ] [tag="PRO.*" | word="ne" ] {0,3} [tag="V.*" ]` ne permettrait pas des phrases telles que « *on ne la voit pas* » car le moteur de recherche ne trouverait que des locutions uniquement avec « *ne* » ou uniquement sans « *ne* » en évitant les possibles combinaisons.

Cette expression régulière n'est toutefois pas exhaustive puisqu'il faut tenir compte de la possibilité d'insérer le « *-t-* » entre le verbe et le pronom *on* en inversion. Ce type d'occurrence reste marginal, pourtant non négligeable (par exemple, dans le sous-corpus littéraire, il apparaît neuf fois dans *Voyage au bout de la nuit* de Céline et six fois dans *Les mots* de Sartre).<sup>8</sup>

## 2.3 Méthodologie de traitement des données

Vu l'étendue du corpus InterCorp, il serait impossible de travailler avec toutes les occurrences du pronom *on* lors de la recherche qualitative. C'est pour cela qu'il faut en choisir quelques-unes en s'appuyant sur une méthodologie bien définie. Les fonctions de l'interface InterCorp qui permettent de créer des échantillons aléatoires seront donc utilisées.

---

<sup>8</sup> Ici, il ne s'agit pas exclusivement des questions formées en inversant le sujet et le verbe. Ces questions ne constituent qu'un quart des occurrences, les trois quarts restants donc représentent les phrases exclamatives, discours indirects ou des phrases stylistiquement marquées.

Ainsi, il est nécessaire d'indiquer la quantité de phrases dans l'échantillon ainsi que la quantité d'échantillons. L'Institut du corpus national tchèque recommande travailler avec au moins trois échantillons dont la taille est égale ou supérieure à 10 % de la totalité des occurrences. Vu que le nombre d'occurrences du pronom *on* est 5760 dans le sous-corpus littéraire et 9571 dans le sous-corpus juridique, cette étude sera basée sur 4 échantillons à 150 phrases (équivalent à 600 occurrences, c'est-à-dire 10,4 %) pour les textes littéraires et 7 échantillons à 150 phrases (équivalent à 1050 occurrences, c'est-à-dire 11 %) pour les textes juridiques.

## 2.4 Ecueils rencontrés

Il faut tenir compte du fait que le corpus InterCorp n'est pas fiable à cent pour cent, surtout en ce qui concerne la langue de source et le tagging (l'étiquetage grammatical). Après définir les conditions de la recherche, le sous-corpus Acquis présente le nombre le plus élevé des mélanges des versions linguistiques et d'autres fautes. Mentionnons au moins quelques types : mélanges avec l'anglais (*on European business*) ou finnois (*on voimassa*) ; fautes typographiques (*on utilis/jait*) ; fautes d'orthographe (*on été modifiées*) fautes d'alignement.<sup>9</sup> Le problème auquel il faut faire face, c'est aussi l'impossibilité de distinguer les verbes lexicaux des verbes auxiliaires.

Les mélanges des variantes linguistiques ne sont pas nécessairement une faute de l'alignement ou d'une mauvaise distinction de la langue source, mais peuvent être aussi citations des langues étrangères, surtout l'anglais. De plus, le verbe est parfois mal conjugué, plus précisément sa conjugaison ne correspond pas à la norme, par exemple « *on peu dire* » au lieu de « *on peut dire* », ce qui influence la recherche. Aussi, les signes de ponctuation font parfois partie du mot, il est donc impossible pour le tagger de distinguer le lemme initial, par exemple « *on pourrait* » (sans espace). Un autre souci pose le discernement de la forme élémentaire, le lemme, du corpus InterCorp qui n'est pas

---

<sup>9</sup> Ce qui rend le travail encore plus ardu, c'est l'alignement trop décalé sans qu'il soit possible de voir le contexte dont nous avons besoin pour l'analyse qualitative. Ce cas n'est pas très fréquent, il rend toutefois l'analyse compliquée puisqu'il faut chercher le texte de source sur Internet.

toujours fiable. Par exemple, le subjonctif du verbe *aller* est *aille*, ce que le corpus comprend comme une forme verbale du verbe *ailler*.<sup>10</sup>

En ce qui concerne le corpus Frantext, à la différence du corpus InterCorp, il n'est malheureusement pas doté d'un moteur de recherche qui permettrait de chercher les différents tags (étiquettes grammaticales). Ainsi, il n'est pas possible de fournir une analyse profonde basée non seulement sur les données du corpus InterCorp, mais aussi sur les données d'un corpus référentiel monolingue.<sup>11</sup>

De plus, Frantext n'est pas capable d'afficher plus de 50 000 résultats, il faut donc effectuer la recherche en plusieurs étapes. C'est pour cela que ce n'est qu'un découpage détaillé (en décennies à partir de 1950) qui permet de comparer les données avec InterCorp en évitant la suppression de plusieurs résultats.

Néanmoins, ce n'est pas que le découpage initial qui rend la recherche plus difficile. Elaborer une liste de fréquences des verbes qui accompagnent le pronom *on* est compliqué aussi car l'étude du voisinage ne permet d'examiner que les 10 000 premiers résultats. De plus, une liste des mots outils exclus des mots en voisinage du pivot est prédéfinie et élimine des verbes tels que *pouvoir* ou *avoir*. Ainsi, un chercheur doit créer sa propre liste des mots outils.

C'est pour toutes ces raisons qu'un travail manuel rigoureux est essentiel pour pouvoir obtenir des résultats fiables et comparables.

## 2.5 Questions de recherche

L'objectif de ce mémoire est d'examiner les différents aspects du pronom *on* surtout quant à son pouvoir de déterminer le référent. Pour ce faire, il est possible de recourir à d'autres langues qui ne disposent pas d'un équivalent direct de ce pronom, par exemple le tchèque et l'espagnol. La question principale est donc la suivante :

---

<sup>10</sup> Dans les sous-corpus sur lesquels est basée cette étude, la forme du subjonctif présent du verbe *aller* se trouve uniquement dans *Voyage au bout de la nuit* de Louis-Ferdinand Céline. Toutes les quatre occurrences sont mal étiquetées.

<sup>11</sup> Même les textes de Frantext contiennent des intrus : citons au moins *father's* (venant de l'anglais), « à » (la faute orthographique – « à » avec l'accent grave au lieu de « a » sans accent) ou « est » (de nouveau une mauvaise graphie).

## **Comment est-il possible d'identifier le degré de détermination du référent du pronom *on* en s'appuyant sur ses équivalents en tchèque et en espagnol ?**

Pour répondre à cette question fondamentale, ce travail cherche progressivement les réponses aux questions partielles suivantes :

- 1) Dans quel type de texte (littéraire ou juridique) se trouvent plus d'occurrences du pronom *on* ?
- 2) Quels sont les verbes les plus utilisés dans le contexte immédiat du pronom *on* et en quoi peuvent-ils influencer le degré de détermination ? Quelle est leur corrélation avec le type de textes et le type d'équivalent ?
- 3) L'accord en nombre et en genre, qui détermine fortement le référent, est-il fréquent ?
- 4) La composition des types d'équivalents tchèques et espagnols du pronom *on* sera-t-elle la même dans les deux types de textes ?
- 5) Les équivalents tchèques et espagnols du pronom *on* disposent-ils de même degré de détermination dans les deux types de textes ?
- 6) En quoi les échelles de la détermination de chaque type de texte sont-elles différentes ?

Vu les différents aspects présentés dans la partie théorique, les hypothèses suivantes vont être confirmées ou infirmées :

- 1) Vu que les textes littéraires peuvent souvent imiter le style oral, supposons que le pronom *on* s'y trouvera plus souvent que dans les textes juridiques.
- 2) Vu que le vocabulaire juridique est assez stéréotypique, il différera davantage du style littéraire et la diversité lexicale influencera la diversité référentielle du pronom *on*.
- 3) Vu qu'un des avantages du pronom *on* est que le référent reste souvent vague, le participe/l'adjectif ne s'accordera pas souvent afin de ne pas déterminer le référent.
- 4) Vu que les procédés formels d'impersonnalisation du tchèque et de l'espagnol sont semblables, la composition des types d'équivalents sera analogue.
- 5) Vu que la diversité référentielle du pronom *on* est élevée, il est probable que le degré de détermination différera car les traducteurs tchèques et espagnols n'interpréteront pas le contexte de la même manière.
- 6) Vu que l'application des textes juridiques est d'habitude générale et non pas explicite, il est probable que les équivalents de ces textes seront dotés d'un degré moins élevé de détermination que les équivalents du pronom *on* dans les textes littéraires.

Tláškal (2005) établit une liste des équivalents tchèques et espagnols qui est présentée dans le chapitre 1.3. Cependant, il ne tient pas compte de la différence systémique entre eux. C'est pour cela que cette étude va essayer de construire une échelle de détermination basée sur les différents types d'équivalents, regroupés en blocs plus généraux. Cette échelle va servir de base pour les analyses qualitatives du corpus.

Afin d'obtenir les réponses aux questions citées ci-dessus, la partie empirique sera divisée en plusieurs chapitres. Dans un premier lieu, les verbes qui accompagnent le pronom *on* seront étudiés afin de pouvoir dégager les tendances générales auxquelles obéissent les textes littéraires et juridiques. Après une étude quantitative des lemmes observés dans les deux sous-corpus, une comparaison entre les verbes les plus fréquents sera effectuée. Ainsi, une base solide pour l'analyse des équivalents pourra être constituée.

Dans un second lieu, les textes littéraires seront affrontés en s'appuyant sur la recherche quantitative des fréquences de chaque type formel des traductions en tchèque et en espagnol. Puis, les exemples-types seront analysés et placés sur l'échelle de la détermination. Enfin, une approximation de placement de la totalité des occurrences sera proposée.

Dans un troisième lieu, la même méthode sera appliquée aux textes juridiques. Finalement, ces deux chapitres seront contrastés afin de pouvoir en dégager une tendance générale applicable sur le pronom *on* en français.

### 3 PARTIE EMPIRIQUE

La troisième partie de ce mémoire est la partie empirique. Dans un premier temps, le pronom *on* sera examiné de point de vue quantitatif : sa fréquence et son voisinage seront analysés. Puis, le fonctionnement du pronom *on* dans les textes littéraires sera contrasté en s'appuyant sur les équivalents tchèques et espagnols. Enfin, la même méthode sera appliquée sur les textes juridiques.

#### 3.1 Pronom *on* du point de vue de fréquence et de son voisinage

Dans les textes littéraires, la fréquence absolue du pronom *on* est 5 760 et la fréquence relative est 804,35 ipm (*instances per million*). Néanmoins, afin de pouvoir dégager les tendances générales, il est nécessaire de comparer les données avec le corpus référentiel de Frantext. Dans l'époque comparable avec celle contenue dans le corpus InterCorp, c'est-à-dire à partir des années 1950, le corpus Frantext dispose de 378 textes contenant le pronom *on* ; le nombre de positions est 33 713 137 et la fréquence absolue 146 044 occurrences. La fréquence relative, le chiffre le plus informatif, est 4 331,96 ipm. Elle est donc cinq fois supérieure à celle observée dans le corpus InterCorp. Bref, il est évident que pour présenter des résultats généralement applicables, une étude basée sur un corpus plus large devrait être effectuée.

Dans les textes juridiques, la fréquence absolue du pronom *on* est 9 571 et la fréquence relative est 306,03 ipm (*instances per million*). Ainsi, il en résulte que la fréquence relative des textes littéraires est presque trois fois supérieure à celle des textes juridiques – le pronom *on* apparaît donc plus souvent dans les textes littéraires. Tout comme dans le cas des textes littéraires, il serait pertinent d'introduire les résultats tirés d'un autre corpus référentiel. Cependant, le corpus Acquis ne peut pas être comparé ni avec le corpus frWac, ni avec Araneum qui n'englobent que les sites Internet. Enfin, il faut tenir compte du fait que dans le cas des exemples tirés des sous-corpus Acquis et Core, l'interprétation des résultats peut être influencée par le nombre de positions limité.



### 3.1.1 Voisinage de la variante *l'on*

Même s'il ne s'agit pas du but primaire de ce travail, afin de mieux comprendre la problématique, il faut s'intéresser non seulement aux fréquences du pronom *on*, mais aussi à son voisinage. D'abord, soulignons que la position prototypique de ce pronom est juste avant le verbe. Néanmoins, d'autres possibilités existent : l'inversion (théoriquement avec l'insertion du « -t- ») ou l'insertion du « l' ».

Le chapitre 1.1.2 a montré que les théoriciens insistent sur la possibilité d'insérer le « l' » avant le pronom *on* si celui-ci suit les conjonctions « *si* », « *et* », « *que* » ou les pronoms relatifs « *que* » et « *où* ». Les résultats tirés du corpus InterCorp (des textes dont le français est la langue source) confirment cette théorie car dans 91,7 % des cas, « *l'on* » est précédé d'une conjonction ou d'un pronom : *que* est le plus fréquent (43,7 %) suivi de *où* (20,7 %), *si* (18,6 %), *et* (8,7 %). Néanmoins, il existe d'autres types de voisinage où le « l' » est inséré pour éviter le hiatus en début de la phrase, par exemple les symboles typographiques : la virgule (1,5 %), guillemets (0,19 % occurrences) ou point (0,14 %). Ces types restent toutefois marginaux.<sup>12</sup>

De plus, il serait pertinent d'examiner aussi la distribution de « *l'on* » dans les textes littéraires et juridiques. L'attention doit être prêtée plutôt à la fréquence relative : dans le sous-corpus Core, la fréquence relative de « *l'on* » est 198,3 ipm (*instances per million*), alors que dans le sous-corpus Acquis se trouvent 64,81 ipm. Ainsi, cette variante est sensiblement plus préférée dans les textes littéraires.

Cependant, pour ce travail, c'est surtout l'accompagnement lexical, ou plus précisément verbal du pronom *on* qui est essentiel puisque la détermination du référent peut aussi être influencée par la nature du verbe qui accompagne le pronom.

Etant donné que cette partie de l'analyse veut montrer la diversité lexicale et les combinaisons du pronom *on* et le lemme verbal, les verbes auxiliaires seront triés et classés dans les cases des verbes à sens plein qui ne sont pas auxiliaires.<sup>13</sup> Bien sûr, comme

---

<sup>12</sup> En plus, il est possible que ce ne soit pas la ponctuation qui engendre l'insertion de « l' », mais plutôt le début de la phrase.

<sup>13</sup> Le logiciel KonText ne permet pas de définir la nature lexicale du verbe. Ainsi, il traite de la même manière les verbes comme *avoir* et *garder* qui, évidemment, sont sémantiquement très différents. Les verbes auxiliaires ne portent pas un sens à proprement parler. Cependant, c'est surtout le sens du verbe qui est essentiel pour ce travail.

le mentionne Blanche-Benveniste (2005, p. 81) : « *les verbes les plus fréquents sont des verbes dont la signification est fondamentale, et semble préexister à celle de tous les autres : avoir, être, aller, faire, pouvoir, voir, devoir, dire, venir* ». Cependant, afin de comprendre si la nature du texte étudié et la sémantique influencent la référence du pronom *on*, il est essentiel que ces verbes ne soient pas classés parmi les auxiliaires puisqu'ils font partie du groupe verbal, mais n'apportent pas le sens. Et c'est surtout le sens qui est intéressant pour cette étude.

D'abord, une étude pour chaque type de texte sera effectuée. Puis, des informations en tirées seront analysées pour mieux illustrer les spécificités, les différences et les similitudes afin de dégager un autre aspect du fonctionnement du pronom *on*.

### **3.1.2 Accord du participe et de l'adjectif avec le sujet du verbe *être***

Dans la totalité du corpus InterCorp, le verbe *être* accompagne le pronom *on* 2286 fois en tant que verbe auxiliaire de la forme du passé composé. Vu que le participe passé s'accorde en nombre et en genre avec le sujet du verbe, cet accord peut éclaircir le référent du pronom *on*, ou au moins le nombre et le genre. Cependant, il faut tenir compte du fait que 76 % de ces occurrences sont tirées du sous-corpus Subtitles dont la fiabilité peut être contestée (surtout à cause de la langue de source et les compétences de chaque traducteur inconnues).<sup>14</sup> En prenant en compte par exemple uniquement les occurrences dans les textes littéraires dont la langue source est le français, c'est-à-dire un sous-corpus fiable, il est possible de dégager la tendance de ne pas accorder le participe avec le sujet réel.

Au total, l'accord en nombre (du participe passé) a été effectué dans 30 % des cas dans les textes littéraires, la fréquence de l'accord des adjectifs est encore moins importante, 12 % s'accordent en nombre.<sup>15</sup> Néanmoins, il faut prêter attention au contexte plus large car

---

<sup>14</sup> Dans certains corpus (comme par exemple Subtitles et Acquis), la langue source n'est pas indiquée. Ainsi, ces sous-titres peuvent être non seulement les originaux français, mais aussi les traductions dont les auteurs ne sont pas repérables. Ce travail ne s'intéresse qu'aux originaux français car la langue de traduction peut interpréter le référent.

<sup>15</sup> Cependant, il est difficile de déterminer si les adjectifs et participes qui finissent par un -x ou un -s sont au pluriel ou au singulier. Considérant qu'ils ne sont pas nombreux et que la tendance générale est plutôt orientée vers le singulier, dans ce travail, ils sont classés parmi les singuliers (donc moins déterminés).

ce chiffre ne signifie pas nécessairement que le référent soit toujours au pluriel. En ce qui concerne les textes juridiques, l'accord n'est jamais mis en place.<sup>16</sup>

Dans de nombreux exemples de l'accord, le pronom *on* est précédé d'un autre pronom personnel déterminant le nombre, surtout *nous* : « **On s'est levés nous, on s'est dégagés de ses mains.** » (Louis-Ferdinand Céline, *Voyage au bout de la nuit*). La capacité de l'accord de dévoiler le nombre de référents peut s'appuyer aussi sur l'observation que dans la même œuvre, les auteurs ont parfois recours à l'emploi du participe accordé et parfois non-accordé, bien qu'il s'agisse du même sens du verbe.<sup>17</sup>

(1) *Deux heures pleines, on est resté ainsi à ne rien prendre, à ne rien faire.* (Louis-Ferdinand Céline, *Voyage au bout de la nuit*)

(2) *Bien fiers alors d'avoir fait sonner ces vérités utiles, on est demeurés là assis, ravis, à regarder les dames du café.* (Louis-Ferdinand Céline, *Voyage au bout de la nuit*)

Les phrases 1 et 2, tirées du *Voyage au bout de la nuit* de Louis-Ferdinand Céline, en font preuve. Dans le premier cas, l'accord est absent, alors que dans la deuxième phrase, quatre accords en nombre sont percevables. Cependant, les verbes lexicaux n'offrent pas cette possibilité de détermination plus ou moins évidente. Le chapitre suivant va donc repérer la variété des verbes qui accompagnent le pronom *on* en fonction du type de texte.

### 3.1.3 Types des verbes en voisinage du pronom *on*

Après avoir vu les fréquences du pronom *on* dans les textes littéraires et juridiques ainsi que l'emploi de la variante *l'on*, il faut maintenant s'intéresser à la nature des verbes qui accompagnent ce pronom. Une analyse comparée des deux types de textes pourra servir de base pour la suite de la partie empirique (surtout les chapitres 3.2 et 3.3).

---

<sup>16</sup> En outre, dans la plupart des cas, le verbe être n'est pas suivi d'un participe, mais plutôt des expressions telles que « être en droit », « être en présence » ou « être loin ».

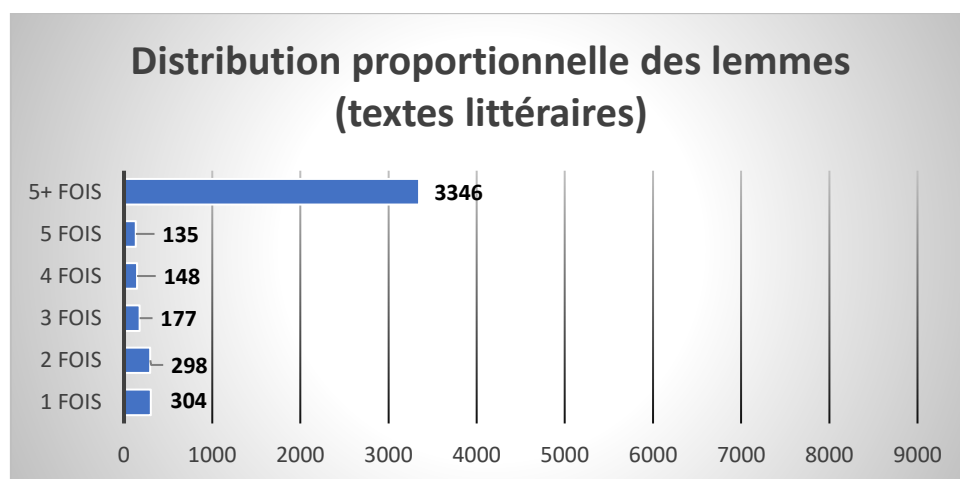
<sup>17</sup> Bien sûr, il faut toujours observer s'il s'agit d'un exemple tiré de la voix du narrateur ou d'un discours d'un personnage. Dans l'exemple 1 et 2, c'est le personnage qui parle, les deux sont donc complètement comparables.

### 3.1.3.1 Verbes accompagnant le pronom *on* dans les textes littéraires

Tout d’abord, cette partie va être concentrée sur le voisinage du pronom *on* dans les textes littéraires afin de fonder une base solide pour l’analyse du degré de détermination. C’est donc aussi le voisinage du pronom *on* et le contexte plus large qui permettent de découvrir la détermination du référent.

Comme l’a déjà évoqué le chapitre 2.5 présentant les hypothèses, contrairement aux textes juridiques, une abondance des lemmes est attendue dans les textes littéraires puisqu’il s’agit d’une des caractéristiques de ce style (voir le chapitre 1.1.1). Le chapitre concentré sur la fréquence de *on* a déjà montré que ce pronom apparaît 5 760 fois. Néanmoins, il est aussi important de s’intéresser à la richesse lexicale des verbes qui l’accompagnent : au total, 690 lemmes figurent dans les textes littéraires.<sup>18</sup>

La figure 1 dépeint la distribution proportionnelle des lemmes et du nombre de leurs occurrences. Par distribution proportionnelle s’entend le produit de la quantité des lemmes différents multipliée par le nombre de répétitions. Ainsi, la fréquence totale des lemmes demeure moins importante que la disparité entre le nombre d’emplois du lemme le plus fréquent et la quantité des lemmes les moins utilisés.



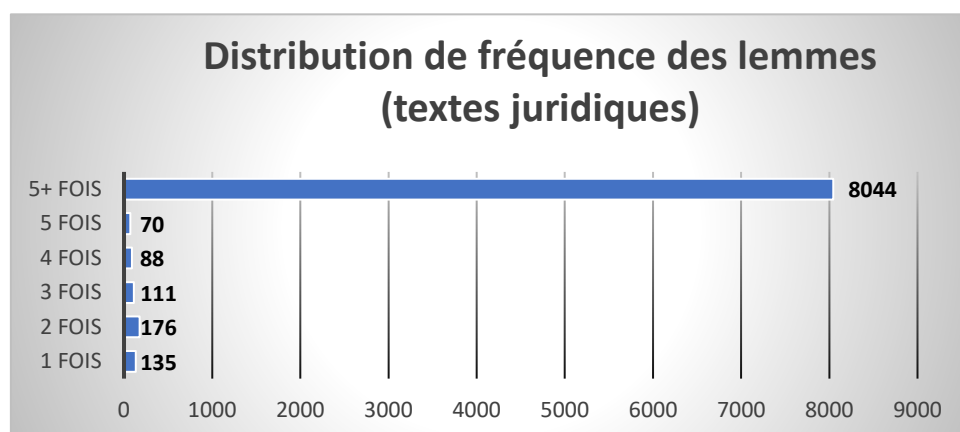
**Figure 1 :** *Distribution proportionnelle des lemmes dans les textes littéraires*

<sup>18</sup> Il serait intéressant de compléter cette analyse avec une étude de voisinage du corpus Frantext pour voir à quel point le corpus InterCorp est influencé par sa composition. Malheureusement, Frantext ne dispose pas de moyens assez pertinents pour une telle analyse car l’interface ne permet pas une recherche aussi avancée. Il est possible de chercher le voisinage du pronom *on*, mais il faut préciser à quelle position par rapport au KWIC le verbe est situé, ce qui rend la recherche très compliquée.

Voici un élément important pour l'analyse du voisinage du pronom *on* : la distribution proportionnelle des lemmes. Cette figure démontre que 76 % des formes verbales contenues dans le sous-corpus littéraire sont des verbes qui apparaissent plus de cinq fois. Ceci confirme que les textes littéraires sont plus abondants que les textes juridiques en ce qui concerne le vocabulaire.<sup>19</sup> Il est intéressant de noter que dans le sous-corpus littéraire, les verbes qui se répètent plus de cinq fois forment 16 % des lemmes, ce qui paraît assez peu. Dans le chapitre suivant, la même étude va être répétée pour les textes juridiques afin de voir à quel point ces deux types de textes sont différents.

### 3.1.3.2 Verbes accompagnant le pronom *on* dans les textes juridiques

Dans le sous-corpus des textes littéraires, 44 % des lemmes n'apparaissent qu'une seule fois. Maintenant, il faut chercher à comparer ces données avec les mêmes observations effectuées dans le sous-corpus des textes juridiques. Ainsi, il sera plus pertinent de décider s'il s'agit d'un chiffre élevé ou non.



**Figure 2 :** *Distribution de fréquence des lemmes dans les textes juridiques*

Le contraste entre la distribution proportionnelle des lemmes dans le sous-corpus littéraire et dans le sous-corpus Acquis est évident grâce à la figure 2. Alors que les verbes qui apparaissent à plus de 5 reprises forment 76 % des occurrences dans les textes littéraires, l'accompagnement lexical du pronom *on* semble beaucoup plus homogène dans les textes juridiques. Seuls 7 % des verbes qui complètent le syntagme verbal sont des verbes qui n'ont plus de 5 occurrences, les 93 % restants sont les répétitions des verbes fréquents. Ainsi, il est possible de constater que les textes juridiques témoignent d'une tendance à limiter la variabilité lexicale plus que les textes littéraires.

<sup>19</sup> Voir le chapitre 1.1.1 pour plus de détails sur les textes juridiques.

### 3.1.3.3 Comparaison des verbes accompagnant le pronom *on* dans les différents types de textes

Cependant, pour mieux illustrer les spécificités de chaque type de texte, une étude comparée des deux sous-corpus et de la totalité d'InterCorp va être présentée. Le tableau ci-dessous montre les lemmes les plus fréquents en termes absolus ainsi que relatifs. Ainsi, une compréhension plus profonde du fonctionnement spécifique des textes littéraires et juridiques peut être atteinte.

Verbe	Totalité d'InterCorp		Textes littéraires		Textes juridiques	
	Fq absolue	Fq relative	Fq absolue	Fq relative	Fq absolue	Fq relative
<i>Avoir</i>	25458	228,81	271	37,88	64	2,08
<i>Pouvoir</i>	21953	197,31	744	104,10	2098	68,18
<i>Aller</i>	10871	97,71	86	12,01	1	0,03
<i>Être</i>	10294	92,52	290	40,66	102	3,32
<i>Faire</i>	6998	62,90	154	21,56	75	2,43
<i>Dire</i>	4543	40,83	161	22,48	7	0,22
<i>Devoir</i>	4331	35,96	94	13,09	208	6,75
<i>Savoir</i>	3588	32,08	62	8,62	124	4,04
<i>Vouloir</i>	3323	29,87	92	12,94	382	12,41
<i>Voir</i>	3346	30,48	110	15,40	57	1,86
<i>Parler</i>	2052	18,44	61	8,47	9	0,29
<i>Venir</i>	1475	13,26	52	7,24	9	0,29
<i>Trouver</i>	1444	12,98	65	9,09	84	2,72
<i>Prendre</i>	1010	9,08	47	6,62	44	1,43
<i>Entendre</i>	681	6,12	47	6,62	3244	105,42
<i>Mettre</i>	409	3,67	43	6,01	37	1,22
<i>Arriver</i>	370	3,32	39	5,39	9	0,29
<i>Finir</i>	286	2,57	47	6,62	0	0,00
<i>Revenir</i>	106	0,95	20	2,77	2	0,08
<i>Constater</i>	589	0,78	1	0,15	165	5,37

**Tableau 4 :** *Les verbes les plus fréquents accompagnant le pronom on*

Le tableau 4 présente les verbes qui accompagnent le plus souvent le pronom *on* dans la totalité du corpus, dans les textes littéraires et dans les textes juridiques. Chaque fois, la fréquence absolue est mise en relation avec la fréquence relative (indiquée en ipm, *instances per million*). Quelques différences entre les fréquences des verbes observés méritent un commentaire plus détaillé. C'est pour cela que le taux de fréquence considérablement moins élevé dans un type de texte que dans l'autre est marqué en bleu clair, le bleu foncé indique

le taux de fréquence excessivement plus élevé. Les données ainsi marquées sont commentées ci-dessous. Voici donc les différences et les similitudes les plus remarquables :

- 1) Le verbe *dire* en combinaison avec le pronom *on* est souvent utilisé dans les textes littéraires, mais très peu fréquent dans les textes juridiques. Ce fait est important pour le chapitre 3.2 dans lequel les équivalents tchèques et espagnols seront examinés. « *On dit que* » est une formule trop vague pour le langage juridique ; le référent du pronom *on* reste totalement indéterminé.
- 2) De même, le verbe *parler* n'est presque pas utilisé dans les textes juridiques, il n'apparaît que 8 fois. Vu qu'il s'agit d'un corpus législatif, il est envisageable que ce verbe ne soit pas souvent présent car ces textes ne contiennent pas des discours indirects. Ainsi, la nécessité présente dans les textes littéraires d'introduire les discours rapportés en employant les verbes de parole est absente dans le sous-corpus Acquis. Cependant, c'est le verbe *constater* dont les fréquences sont beaucoup plus élevées dans les textes juridiques que dans les textes littéraires. De nouveau, c'est un verbe assez typique des textes législatifs et des discours prononcés par des politiques.
- 3) Alors que les verbes d'état (par exemple *être* ou *devenir*) n'apparaissent que très peu dans les textes juridiques, ils sont beaucoup plus présents dans les textes littéraires. Ceci peut être expliqué par l'observation que les verbes auxiliaires sont presque effacés des textes de l'Union européenne car le temps prépondérant est le présent atemporel, le présent des vérités générales. Néanmoins, il est possible de constater que dans la totalité du corpus, la fréquence relative du verbe *être* semble encore plus élevée que dans les textes littéraires.
- 4) Le verbe *venir* n'est presque pas utilisé dans les textes juridiques, seules 8 occurrences en témoignent. De plus, il s'agit exclusivement de son usage semi-auxiliaire, il indique que l'action exprimée par l'infinitif s'est achevée très récemment : « *Dans le contexte de l'observation que l'on vient de formuler [...].* »
- 5) Le verbe qui accompagne le pronom *on* le plus fréquemment dans les textes juridiques est le verbe entendre. Ceci peut être expliqué par sa fonction sémantique : il introduit les définitions qui sont tellement importantes dans les documents législatifs. Le verbe

*définir* en lui-même n'apparaît que 7 fois, une tendance à utiliser la locution figée « *on entend par* » semble donc être préférée par le langage de l'Union européenne. Cette locution est très typique pour le langage législatif puisqu'elle s'emploie pour introduire les définitions dans de nombreux actes : les résolutions, les directives, les règlements, etc. Le langage de l'Union européenne peut présenter des spécificités non seulement au niveau lexical, mais aussi syntaxique. La collocation « *on entend par* » est donc presque figée dans ce type de textes.

- 6) Après avoir mentionné le verbe *venir* prenant fonction du semi-auxiliaire, il faut aussi prendre en considération que le verbe *aller* oscille entre la catégorie des semi-auxiliaires et des verbes lexicaux (à sens plein). Cependant, il n'apparaît qu'une seule fois dans le sous-corpus des textes juridiques. L'explication semble assez évidente : le temps verbal prédominant est le présent atemporel, parfois aussi le futur atemporel. Néanmoins, le futur proche formé par le verbe *aller* + infinitif n'est pas typique pour ce type de texte.

Pour résumer cette analyse, il est essentiel de souligner que le verbe *entendre* domine dans l'accompagnement lexical du pronom *on* dans les textes juridiques. Cette observation demeure décisive lors de la traduction, ce qui va être démontré dans le chapitre 3.2. De l'autre côté, dans les textes littéraires ne prédomine aucun verbe excepté *pouvoir* (dont le sens modal et pas pleinement lexical doit être souligné).

De plus, 93 % des verbes contenus dans le sous-corpus Acquis sont des verbes qui se répètent plus de 5 fois, il est donc concevable que la richesse formelle des équivalents tchèques et espagnols soit inférieure à la diversité des équivalents du pronom *on* dans les textes littéraires. La répartition des lemmes dans les textes littéraires semble beaucoup plus proportionnée que dans les textes juridiques, il est donc possible de s'attendre à une abondance des équivalents littéraires variés, contrairement à une certaine uniformité juridique.

Au vu de ces constatations, il est probable que l'échelle de la détermination du référent du pronom *on* dans les textes juridiques (telle que définie dans le chapitre 3.2.3) ne sera pas proportionnellement représentée, tandis que l'échelle dans les textes littéraires (telle que définie dans le chapitre 3.2.2) sera plus proportionnellement englobée.



### 3.2 Analyse des équivalents dans les textes littéraires

Dans ce chapitre, une analyse des équivalents du pronom *on* dans les textes littéraires sera effectuée. D'abord, la composition des équivalents regroupés en 10 catégories sera présentée. Puis, les correspondances entre les équivalents seront examinées du point de vue quantitatif, mais aussi qualitatif. Finalement, ces équivalents seront placés sur l'échelle de la détermination et comparés du point de vue du degré de détermination.

	Tchèque			Espagnol		
	Rang	Fq absolue	% dans l'échantillon	Rang	Fq absolue	% dans l'échantillon
1 <sup>ère</sup> personne du pluriel	1	137	22,8 %	1	145	24,2 %
3 <sup>e</sup> personne du pluriel	2	96	16,0 %	2	143	23,8 %
<i>Člověk/lidé/gente</i>	3	93	15,5 %	9	2	0,4 %
Passif réfléchi	5	72	12,0 %	3	142	23,6 %
Construction impersonnelle ( <i>lze</i> etc.)	6	37	6,2 %	6	26	4,4 %
Équivalent zéro	7	28	4,7 %	7	16	2,7 %
Passif promotionnel	8	24	4,0 %	10	0	0,0 %
<i>Někdo/nikdo/alguno/uno/nadie</i>	9	23	3,8 %	8	27	4,5 %
2 <sup>e</sup> personne du pluriel/singulier	10	17	2,9 %	4	52	8,7 %
Autres	4	73	12,1 %	5	46	7,7 %

**Tableau 5 :** Composition des équivalents du pronom *on* en tchèque et en espagnol (dans les textes littéraires)

Ce tableau présente dix catégories d'équivalents observées dans le sous-corpus des textes littéraires, classées selon le taux de leur fréquence dans l'échantillon de 600 occurrences (10,4 % de la totalité du sous-corpus). La première colonne indique le nom de la catégorie, la deuxième marque le rang de la fréquence de cet équivalent en tchèque, la troisième montre la fréquence absolue des occurrences dans l'échantillon, la quatrième indique le taux de fréquence en tchèque.<sup>20</sup> De même, la cinquième, la sixième et la septième colonne

<sup>20</sup> Afin de démontrer les tendances de l'emploi des équivalents du pronom *on* dans les deux langues de la manière la plus claire possible, le classement est organisé selon les résultats observés en tchèque avec la catégorie *Autres* en dernière position bien qu'elle ne soit pas la moins représentée.

représentent le rang de la fréquence, la fréquence absolue et le taux de fréquence de chaque équivalent en espagnol.

Les équivalents dont le taux de fréquence équivaut ( $\pm$  deux points de pourcentage) sont marqués en jaune. Observons qu'il s'agit de la première personne du pluriel, la catégorie la plus représentée dans les deux langues, les constructions impersonnelles, les équivalents zéro et les pronoms indéfinis. De plus, trois de ces quatre types d'équivalents occupent la même position dans le classement selon la fréquence. Les équivalents dont le taux de fréquence est supérieur que leurs homologues dans l'autre langue sont marqués en rouge ; le contraire, c'est-à-dire les équivalents sensiblement moins représentés, en vert.

Il est intéressant que le passif promotionnel n'apparaisse jamais en espagnol en fonction d'équivalent du pronom *on* bien qu'il soit utilisé assez fréquemment comme moyen d'impersonnalisation. Ceci peut être dû non seulement à la tendance des traductions des textes littéraires, mais aussi à l'étendue insuffisante du sous-corpus ou à la composition spécifique des textes.<sup>21</sup>

Alors que les occurrences des noms *člověk* et *lidé* en tchèque sont nombreuses, l'équivalent *la gente* n'est presque jamais présent.<sup>22</sup> En revanche, le tchèque n'a pas souvent recours à l'emploi de la deuxième personne du singulier ou du pluriel et les occurrences du passif réfléchi et de la troisième personne du pluriel sont considérablement moins fréquentes en tchèque qu'en espagnol. Cette observation est digne d'attention car les théoriciens de ces deux langues indiquent que la troisième personne du pluriel et le passif réfléchi sont employés dans des situations semblables. C'est pour cela qu'il est intéressant que l'espagnol prédomine complètement dans les deux catégories. Ainsi, il serait pertinent de comparer quand et dans quelles catégories les équivalents tchèques et espagnols se correspondent.

---

<sup>21</sup> Pour mieux comprendre cette problématique, il convient voir le chapitre 3.3.1 qui dresse un tableau similaire pour les textes juridiques.

<sup>22</sup> L'expression *la gente* en fonction d'un pronom indéfini est plutôt favorisée à l'oral, ceci ne veut pourtant pas dire que les pronoms indéfinis soient totalement absents à l'écrit. Pourtant, à la différence du tchèque, l'espagnol préfère l'emploi du pronom *uno* pour exprimer le degré bas de la détermination. Cette idée va être développée dans le chapitre traitant l'échelle de la détermination.

### 3.2.1 Correspondances entre les équivalents du pronom *on* en tchèque et en espagnol dans les textes littéraires

Après avoir analysé et comparé la composition des équivalents en tchèque et en espagnol, il faut maintenant prêter attention à la correspondance entre ces équivalents. C'est pour cela que deux tableaux vont être constitués afin de démontrer dans quelles catégories les langues se correspondent et dans quelles catégories elles ne trouvent aucune intersection.

	Moyenne dans l'échantillon	Proportion de la totalité
1 <sup>ère</sup> personne du pluriel	24,75	16,50 %
3 <sup>e</sup> personne du pluriel	20,25	13,50 %
Pronom indéfini	4	2,67 %
Passif réfléchi	7,25	4,83 %
Construction impersonnelle ( <i>lze</i> etc.)	0	0,00 %
Équivalent zéro	1,25	0,83 %
Passif promotionnel	0	0,00 %
2 <sup>e</sup> personne du pluriel/singulier	3	2,00 %
Correspondances au total	60,75	40,50 %

**Tableau 6 :** *Correspondance entre les équivalents – proportion de la totalité de l'échantillon (dans les textes littéraires)*

Voici le tableau qui décrit le nombre moyen de correspondances entre les équivalents tchèques et espagnols dans quatre échantillons des textes littéraires dont l'étendue est 150 occurrences. La première colonne dénote le type d'équivalent. Dans la deuxième colonne, la moyenne de nombre de correspondances dans l'échantillon est marquée. La troisième colonne indique la proportion de ces correspondances dans la totalité de l'échantillon.

Remarquons que le nombre total des correspondances est 40,5 %, ce qui est assez étonnant vu la diversité des équivalents dans les deux langues et vu le caractère problématique d'utilisation de deux procédés dont le degré de détermination soit semblable (c'est-à-dire l'interprétation du référent). Il est toutefois nécessaire d'avouer que 30 % en sont les catégories morphosyntaxiques de la première et troisième personne du pluriel (dont le degré de détermination contestable va être commenté plus bas). Au contraire, le passif

promotionnel et les constructions impersonnelles ne se correspondent jamais, ce qui est surprenant aussi car ils entrent dans le même degré de détermination. L'observation de ce degré inégal donc semble cruciale et sera plus développée dans la suite de ce chapitre.

	Nº d'occurrences (réunion tchèque + espagnol)	Nº de correspondances (intersection)	Quotient
<b>1<sup>ère</sup> personne du pluriel</b>	172	92	54 %
<b>3<sup>e</sup> personne du pluriel</b>	147	76	52 %
<b>Pronom indéfini</b>	121	15	12 %
<b>Passif réfléchi</b>	172	27	16 %
<b>Construction impersonnelle (<i>lze</i> etc.)</b>	60	0	0 %
<b>Équivalent zéro</b>	37	5	13 %
<b>Passif promotionnel</b>	22	0	0 %
<b>2<sup>e</sup> personne du pluriel/singulier</b>	54	11	21 %

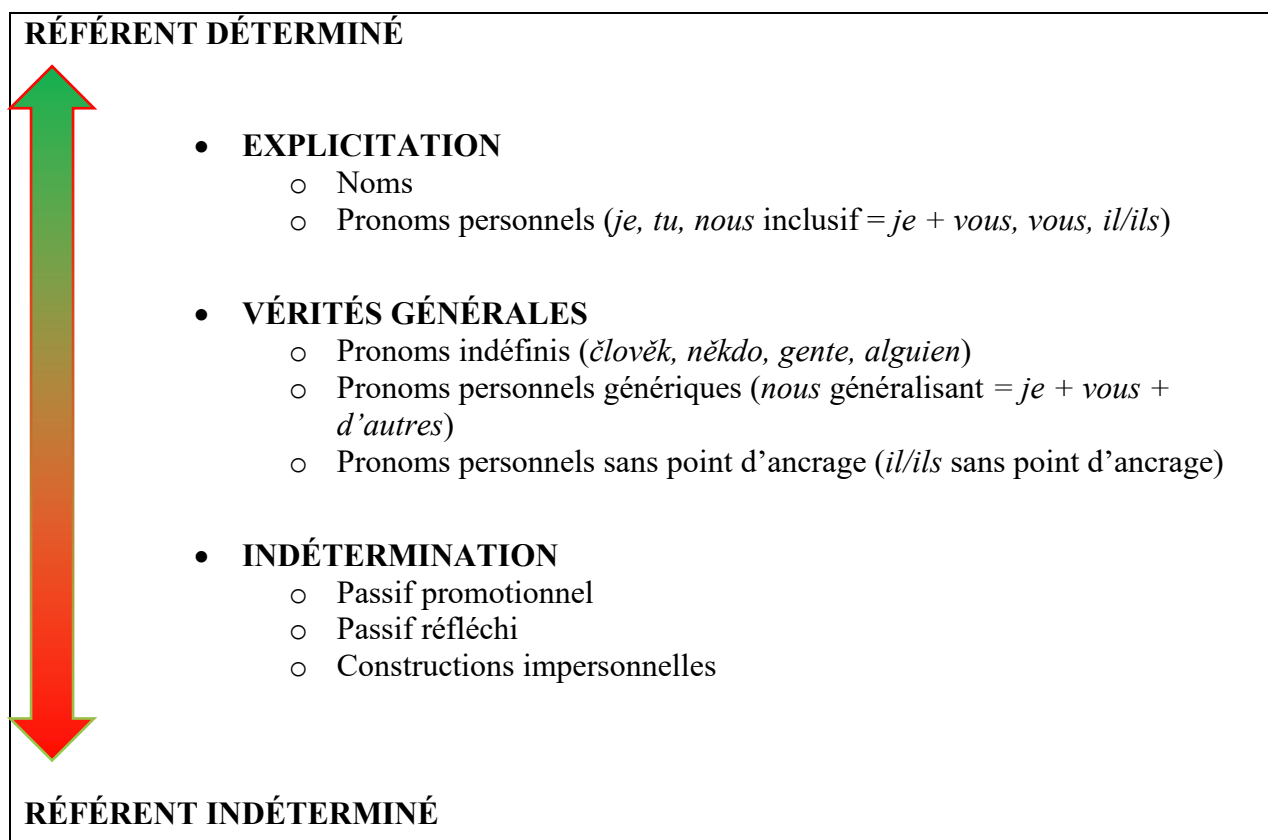
**Tableau 7 :** *Correspondance des catégories d'équivalents en tchèque et en espagnol (dans les textes littéraires)*

Ce tableau montre d'abord le nombre d'occurrences de chaque catégorie dans les deux langues, puis le nombre d'occurrences dans lesquelles la catégorie d'équivalent en espagnol et en tchèque se correspondent. La quatrième colonne est le résultat de la division (en forme de pourcentage arrondi) où le numérateur est la troisième colonne (c'est-à-dire l'intersection des correspondances des équivalents tchèques et espagnols) et le dénominateur est la deuxième colonne (c'est-à-dire la réunion desdits équivalents).

Elle révèle que le tchèque et l'espagnol ont le plus souvent recours à la même catégorie dans le cas de la première et la troisième personne du pluriel. En revanche, les équivalents ne se correspondent jamais dans les catégories du passif promotionnel ou des constructions impersonnelles. Voyons donc l'échelle de la détermination qui va mieux dégager la différence entre le degré de détermination des catégories qui se correspondent le plus et celles qui ne se correspondent jamais.

### 3.2.2 Echelle de la détermination dans les textes littéraires

Les équivalents peuvent être regroupés en 3 degrés de détermination percevables sur l'échelle de la détermination. En vert, voici les types les plus explicites. Le passage de transition entre le vert et le rouge signifie les équivalents qui dénotent des vérités générales applicables sur tous, dont le référent peut être aussi bien déterminé qu'indéterminé. En bas de l'échelle, en rouge, voici les moyens d'expression de l'indétermination totale.



**Figure 3 :** L'échelle de la détermination (dans les textes littéraires)

Deux remarques importantes doivent être faites.

(1) Dans le cadre d'un degré de détermination, une certaine hiérarchie peut être établie aussi. Cette hiérarchie est représentée de manière verticale : les noms sont plus explicites que les pronoms personnels, les pronoms indéfinis (tels que *člověk, la gente, někdo, alguien*) tendent plus vers un degré plus haut de détermination que les pronoms personnels sans point d'ancrage. De l'autre côté, il serait très difficile de décider quel est l'équivalent le moins déterminé, le degré de détermination le plus bas est donc une exception : les trois catégories rangées dedans sont plus ou moins homologues.

(2) Il existe aussi de nombreux exemples à la limite entre les degrés. Par exemple, *uno* est classé parmi les vérités générales car il est applicable sur tous les êtres humains, pourtant caressant d'indétermination totale. Dans un certain contexte, il peut toutefois être compris d'une façon tout à fait explicite (voir la note en bas de page sous l'exemple 5). Comme tout au long de cette étude, c'est ici aussi qu'il faut souligner l'impact énorme du contexte qui permet de mieux évaluer le degré de détermination car il n'est pas possible de se fier aveuglément aux règles établies par l'échelle.

### 3.2.2.1 Explicitation

Bien que le pronom *on* puisse rendre le référent moins évident, il est parfois possible de le déduire à partir du contexte. Cette information est essentielle : c'est surtout le contexte qu'il faut regarder. De plus, il n'est pas question que du contexte immédiat (par exemple KWIC  $\pm 10$ ) – son étendue peut être beaucoup plus large. Ainsi, le contexte aide la détermination. De l'autre côté, il peut mener aux interprétations erronées aussi.

#### I) Noms

La première catégorie du degré le plus haut de détermination est l'explicitation par l'usage des noms. Ici, il faut toujours observer le contexte car l'explicitation peut être basée sur un référent déjà évoqué. Voyons donc deux exemples qui servent à montrer comment ce procédé fonctionne, puis à révéler que parfois, une explicitation qui va trop loin peut causer un glissement de sens important.

(1) *J'étais titularisé, **on avait eu** la bonté de me donner un avenir et je le proclamais enchanteur mais, sournoisement, je l'abominais.*<sup>23</sup> (Jean-Paul Sartre, *Les Mots*)

- *Obdržel jsem titul, **rodina** mi laskavě dala budoucnost, prohlašoval jsem, že je okouzlující, potměšile jsem ji však proklínal.* (traduit par Dagmar Steinová)

La traduction tchèque est l'exemple par excellence de l'explicitation contextuelle. L'information sur la nature de *on* n'est pas contenue dans la phrase source ni dans les phrases

---

<sup>23</sup> Cependant, la traduction espagnole ne tient pas compte de cette subtilité surgie du contexte : « *Estaba titularizado, **habían tenido** la bondad de darme un porvenir y lo proclamaban encantador, pero, disimuladamente, lo aborrecía* » (traduit par Manuel Lamana). L'espagnol traduit le pronom *on* par la troisième personne du pluriel et même offre une interprétation erronée de la partie « *je le proclamais enchanteur* » par « *lo proclamaban encantador* » (au lieu de *je*, le sujet est *ils*).

qui l'entourent, mais beaucoup plus avant quand le personnage parle de son histoire. Ainsi, le pronom *on* devient *la famille* en tchèque.

Néanmoins, n'oublions pas que remplacer le pronom *on* par un nom peut trop restreindre l'éventail des équivalents potentiels. L'exemple suivant fait preuve du fait que parfois, même le contexte plus large ne suffit pas pour déterminer le référent du pronom *on* avec certitude.

(2) *Le directeur lui avait dit qu'il comprenait très bien, mais qu'on ne pouvait pas faire d'exception, qu'il allait voir, mais qu'en somme la situation était grave et que l'on ne pouvait rien décider.*<sup>24</sup> (Albert Camus, *La Peste*)

- *Ředitel mu odpověděl, že to vše dobře chápe, ale že nemůže dělat výjimky, že se mu vynasnaží pomoci, ale že situace je vážná a že on sám nemůže o ničem rozhodovat.* (traduit par Milena Tomášková)
- *El director le había respondido que lo comprendía muy bien, pero que no podía hacer excepciones, que vería, pero que, en suma, la situación era grave y que no se podía decidir nada.* (traduit par Rosa Chacel)

Ici, la même tendance est observable dans les deux langues : celle de déterminer le sujet réel, c'est-à-dire *le directeur*. Néanmoins, le référent du pronom *on* est volontairement flou en français, le tchèque et l'espagnol vont alors très loin dans l'interprétation. Aussi, il est intéressant que le tchèque et l'espagnol explicitent le référent dans la première occurrence du pronom *on* dans cette phrase, mais alors que le tchèque suit cette interprétation explicite, l'espagnol renvoie à l'implicite dans la deuxième occurrence (le passif réfléchi).

Finalement, il serait bien sûr possible de classer cet exemple parmi les pronoms personnels anaphoriques. Vu que la référence est explicitée par un nom, il trouve sa place ici car ce qui est le plus important, ce n'est pas que la catégorie précise, c'est surtout le degré de détermination qui reste le même.

---

<sup>24</sup> En renvoyant au chapitre 1.1.2 sur l'insertion de « l' », il faut souligner que l'insertion observée dans cet exemple est provoquée par la volonté d'éviter le hiatus. Cependant, une tendance intéressante peut être perçue : malgré le contexte immédiat identique, la première occurrence manque cette insertion, elle n'est employée qu'avant la deuxième occurrence du pronom *on*. L'interprétation de ce choix linguistique est difficile à concevoir : c'est probablement pour marquer un détachement encore plus important que dans la première occurrence. Notons aussi que la traduction tchèque va dans le sens opposé : vers une explicitation encore plus soulignée *on sám*.

## II) Pronoms personnels (*je, tu, nous inclusif = je + vous, vous, il/ils*)

Prenant en considération la perspective de l'échelle de la détermination, un des éléments les plus problématiques à placer sur cette échelle sont les pronoms personnels de la première et la troisième personne du pluriel. Ces personnes peuvent être anaphoriques, dont classables parmi les procédés d'explicitation, ou bien généralisantes, dont classables parmi les vérités générales. Il ne faut pas cesser de souligner que c'est toujours le contexte qui aide à fixer le degré de détermination.

(3) *Grouillez-vous de boire ça, et puis **on va** danser.* (Boris Vian, *J'irai cracher sur vos tombes*)

- *Hod'te to do sebe a pak si **půjdeme** zatancovat.* (traduit par Martin Řízek et Petr Turek)
- *Terminatelo de prisa y **nos vamos** a bailar.* (traduit par Jordi Martí Garcés)

Ainsi, l'exemple 3 est un exemple prototypique qui présente une explicitation basée sur le contexte immédiat – il est évident que l'ensemble dénoté par la première personne du pluriel est *je* et *vous*, mais personne d'autre. Cependant, surtout dans le cas de la première personne du pluriel, il faut prêter attention à l'interprétation des rapports externes à la phrase.

Notons aussi que les équivalents tchèque et espagnol se correspondent. Or, comme le chapitre 3.2.1 sur les correspondances entre les équivalents l'a déjà indiqué, cette catégorie dispose du quotient de correspondance le plus élevé.

(4) *À les voir s'éparpiller, gesticuler minutieusement de la sorte et se perdre en dentelles de mouvements saccadés et follement inutiles, **on en demeurait** découragé jusqu'au marasme.* (Louis-Ferdinand Céline, *Voyage au bout de la nuit*)

- *Když jsem je tak viděl běhat, přesně a svědomitě gestikulovat a mořit se s girlandami různých a bláznivě zbytečných pohybů, **býval jsem** z toho na smrt sklíčený.* (traduit par Jaroslav Zaorálek)

Finalement, le dernier exemple de degré de détermination le plus élevé, l'exemple 4, montre un référent trop explicité car ni le contexte ne suggère que le sujet soit explicitement « *je* ». Au contraire, il semble que cette phrase renvoie plus au prochain degré, celle de la vérité générale. La solution du traducteur tchèque fait donc glisser le sens de l'énoncé en le faisant perdre la valeur générale.



### 3.2.2.2 Vérités générales

Le degré intermédiaire entre l'explicitation et l'indétermination est représenté par les vérités générales. Etant donné que les exemples de ce groupe sont applicables sur l'ensemble de l'humanité (car le référent du pronom *on* doit toujours être animé, voir le chapitre 1.1), la première catégorie qui remplit ces conditions sont des pronoms indéfinis (par exemple *člověk, někdo, gente, uno*). Puis, les pronoms personnels génériques et sans point d'ancrage vont être étudiés car contrairement à leurs homologues avec dénotation explicite, ils élargissent la référence à un ensemble général des personnes.

#### I) Pronoms indéfinis (*člověk, někdo, gente, uno*)

La première possibilité d'expression des vérités générales sont les pronoms indéfinis. Bien que le procédé tchèque *člověk* soit formellement classé parmi les noms, il est utilisé en fonction d'un pronom indéfini exprimant une valeur générale. C'est pour cela qu'il fait partie de cette catégorie. De plus, c'est surtout en tchèque que les pronoms indéfinis sont souvent employés en tant qu'équivalents du pronom *on*. De l'autre côté, ce n'est que dans de rares cas que l'espagnol et le tchèque ont recours à cette catégorie dans une même phrase. L'un de ces exemples est le suivant :

(5) *Si on va doucement, on risque une insolation.* (Albert Camus, *L'Étranger*)

- *Když jde člověk pomalu, lehko si užene úpal.* (traduit par Miloslav Žilina)
- *Si uno anda despacio, corre el riesgo de una insolación.* (traduit par José Ángel Valente)

L'exemple 5 est un cas d'impersonnalisation sémantique par le nom ou bien pronom indéfini tchèque *člověk* et le pronom espagnol *uno*.<sup>25</sup> Le plus souvent, l'équivalent tchèque *člověk* correspond à la deuxième personne du singulier en espagnol. Ceci peut être expliqué aussi par la différence de l'emploi du tutoiement car en espagnol, il n'est pas si strictement réservé au milieu familial, son usage est plus fréquent. Il sera intéressant de voir si cette catégorie assez fréquemment présente dans les textes littéraires trouvera le même nombre d'occurrences dans les textes juridiques.

---

<sup>25</sup> Gómez Torrego (1994, p. 14–15) suggère que le pronom *uno* peut non seulement exprimer une vérité générale, mais il peut aussi couvrir un référent concret, un « je » déterminé et dilué en vérité générale. « *¡Qué bien vives, tío! Es que uno sabe montárselo.* » Dans ce cas, il est possible de déduire que *uno* signifie *je*. Néanmoins, c'est surtout le contexte qui permet de décider quel degré de détermination est en question.

## II) Pronoms personnels génériques (*nous* généralisant = *je* + *vous* + *d'autres*)

Les pronoms personnels ont déjà été vus lors de l'analyse des procédés explicitant le référent. Ici, tout comme dans les exemples précédents, il faut recourir au contexte pour voir quel degré de détermination est en question. De plus, la décision sur la généralité/l'anaphore change aussi le placement de l'exemple sur l'échelle de la détermination.

(6) Dans les cas de décès, **on a** toujours tendance à dire une connerie de ce genre. (Michel Houellebecq, *Les Particules élémentaires*)

- *V případě úmrtí máme vždy sklon k podobným žvástům.* (traduit par Alan Beguivin)

Dans cet exemple, c'est aussi la valeur sémantique et l'accompagnement lexical qui justifient le classement parmi les vérités générales. Remarquons la présence de l'adverbe *toujours* (*vždy*) et du nom *tendance* (*sklon*). Néanmoins, bien que l'exemple 6 soit presque sans doute classable comme une vérité générale, il faut réfléchir sur chaque cas avec prudence. C'est une des catégories à la limite entre les degrés de détermination plus haut et moyen.

## III) Pronoms personnels sans point d'ancrage (*il/ils* sans point d'ancrage)

Le même constat que dans le cas du *nous* généralisant peut être fait dans le cas des pronoms personnels *il/ils* qui n'ont aucun point d'ancrage. Bien sûr, il faut toujours regarder le contexte (parfois un contexte très large) qui peut dévoiler le point d'ancrage. Néanmoins, le tchèque et l'espagnol étant des langues qui n'expriment pas obligatoirement le sujet, la présence de cette ancre n'est pas nécessaire. En ce qui concerne le rang de fréquence des occurrences, cet équivalent est placé à la deuxième position.

(7) Les Henrouille eux, n'en revenaient pas d'avoir passé à travers la vie rien que pour avoir une maison et comme des gens qu'**on vient de désemurer** ça les étonnait. (Louis-Ferdinand Céline, *Voyage au bout de la nuit*)

- *Henrouillovi nevyházeli z údivu nad tím, že prožili celý život jenom proto, aby měli vlastní dům; udivovalo je to jako lidi, co je zrovna vytáhli ze sklepa, kde byli zazděni.* (traduit par Jaroslav Zaorálek)
- *Los Henrouille, en cambio, no salían de su asombro por haber pasado por la vida nada más que para tener una casa e, igual que las personas a las que **acaban de sacar** de un encierro entre cuatro paredes, les resultaba extraño.* (traduit par Carlos Manzano)

En guise du dernier exemple des vérités générales, l'exemple 7, il faut introduire l'emploi du pronom *ils* sans point d'ancrage. Ainsi, ce n'est même pas le contexte large qui puisse éclairer le référent. Cet usage est très typique pour le tchèque tout comme pour l'espagnol. Il s'agit donc de l'exemple le moins déterminé des vérités générales car il s'approche beaucoup aux cas d'indétermination.

### 3.2.2.3 Indétermination

Le troisième degré sur l'échelle de la détermination est celui de l'indétermination. Il est représenté par deux types du passif (le passif promotionnel et le passif réfléchi) et par les constructions impersonnelles. Bien sûr, d'autres exemples qui ne rentrent pas dans ces catégories peuvent exprimer le degré d'indétermination. Néanmoins, examinons d'abord les exemples prototypiques.

#### I) Passif promotionnel

La voix active implique souvent le référent alors que la voix passive peut l'éliminer ou au moins le réduire.<sup>26</sup> En tchèque de même qu'en espagnol, il y a deux types du passif : le passif promotionnel (formé par le verbe *být/ser* et le participe passé) et le passif réfléchi (formé par un verbe et le pronom réfléchi *se*).

(8) *Entre Corneille et Pardaillan, on me somrait de choisir.* (Jean-Paul Sartre, *Les Mots*)

- *Byl jsem vyžván, abych si vybral mezi Corneillem a Pardaillanem.* (traduit par Dagmar Steinová)

L'exemple 8 constitue un exemple prototypique du passif promotionnel employé en tant qu'équivalent du pronom *on*. Il a déjà été évoqué que ses occurrences sont assez limitées en espagnol par rapport au tchèque. C'est pour cela qu'il faut se demander quels sont les autres moyens qui peuvent le remplacer et si leur degré de détermination reste le même.

Notons donc que dans les textes littéraires, le passif promotionnel n'est jamais utilisé comme équivalent en tchèque et en espagnol en même temps. D'habitude, quand le passif promotionnel est employé en tchèque, c'est le passif réfléchi auquel l'espagnol a recours,

---

<sup>26</sup> Cela ne veut pas dire que la voix passive exprimée par le passif promotionnel ne puisse jamais expliciter l'agent, bien au contraire. Cependant, pour ce faire, il faut insérer un groupe prépositionnel (en français généralement *par* + agent, en espagnol généralement *por* + agent) ou un groupe nominal à l'instrumental (en tchèque).

le degré de détermination reste alors le même. Néanmoins, la troisième personne du pluriel dans son usage généralisant, qui exprime un degré plus élevé de détermination (la vérité générale), réapparaît fréquemment aussi en tant qu'équivalent du pronom *on* dans les phrases où le tchèque emploie le passif promotionnel. Il est donc évident que l'interprétation du degré de détermination n'est pas nécessairement manifeste.

## II) Passif réfléchi

Dans l'échantillon des textes littéraires, le passif réfléchi est le troisième équivalent du pronom *on* le plus utilisé en espagnol (23,6 % des occurrences)<sup>27</sup>. Il est toutefois sensiblement moins représenté en tchèque (12 % des occurrences). Comme le suggèrent les grammairiens espagnols (voir le chapitre 1.3.1 sur les équivalents du pronom *on* en espagnol), les phrases dont le prédicat est en forme du passif réfléchi sont syntaxiquement impersonnelles car leur sujet est zéro. Ainsi, ce sont les exemples les plus prototypiques de l'indétermination.

(9) *Ils veulent avoir lu le livre recommandé par leur club, celui dont **on** parle, et ils se moquent bien de ce qu'il y a dedans.* (Boris Vian, *J'irai cracher sur vos tombes*)

- *Lidé si chtějí hlavně přečíst knihu doporučenou jejich klubem, knihu, o níž **se** mluví, a kašlou na to, co je uvnitř.* (traduit par Martin Řízek et Petr Turek)
- *El libro que quieren leer es el que recomienda su club, el libro del que **se** habla, y su contenido les importa un bledo.* (traduit par Jordi Martí Garcés)

Vu que le passif réfléchi est beaucoup plus présent parmi les équivalents espagnols que parmi les tchèques, il faut observer quels sont les procédés qu'emploie le tchèque à son lieu. Il s'agit des constructions impersonnelles telles que *je možné* ou *lze* qui disposent d'un même degré de détermination. Néanmoins, la plupart des équivalents sont des vérités générales (les pronoms personnels *nous* et *ils* dans leur emploi généralisant), un déplacement sur l'échelle de la détermination a donc lieu.

---

<sup>27</sup> Cette fréquence du passif réfléchi élevée peut être expliquée aussi par le voisinage du pronom *on*. Par exemple, en espagnol, dans 48 % des cas, la locution « on dit » se traduit par « *se dice* ». De même, 46 % des traductions du verbe *pouvoir* sont classées parmi le passif réfléchi. Voir le chapitre 3.1.3 pour plus de précisions sur les verbes qui accompagnent le pronom *on*.

### III) Constructions impersonnelles

La dernière catégorie qui fait partie du degré de détermination le plus bas est représentée par les constructions impersonnelles. La palette des expressions qui sont classées dans cette catégorie est assez vaste même s'il ne s'agit que de la sixième catégorie la plus présente dans l'échantillon des textes littéraires.

(10) *Dans le silence qui suivit cette démonstration vocale, **on entendit** nettement une mouche traverser l'atmosphère de la pièce avant de se poser sur le visage de Jane.* (Michel Houellebecq, *Les Particules élémentaires*)

- *V tichu následujícím po této hlasové ukázce **bylo** zřetelně **slyšet**, jak pokojem proletěla moucha a usedla na Janinu tvář.* (traduit par Alan Beguivin)

En tchèque, ce sont surtout les verbes de perception (comme par exemple *vidět*, *slyšet*) qui peuvent prendre la forme d'une construction à sujet nul. Comme il a déjà été évoqué, un classement indiscutable est irréalisable. Les constructions à sujet nul peuvent être aussi bien être rangées parmi les vérités générales, par exemple : *Z jejího vrcholu je vidět na jedné straně pláž zvolna klesající k moři [...]* (Michel Houellebecq, *Les Particules élémentaires*, traduit par Alan Beguivin). Bien sûr, le référent n'est pas explicite, cependant, étant donné qu'il s'agit d'un fait atemporel, il serait possible de l'interpréter comme une vérité générale. C'est donc le contexte qui peut aider à identifier le degré de détermination du référent.

(11) *Quand le mystère est trop impressionnant, **on n'ose pas** désobéir.* (Antoine de Saint-Exupéry, *Le Petit prince*)

- *Cuando el misterio es demasiado impresionante, **es imposible** desobedecer.* (traduit par Bonifacio del Carril)

En espagnol, ce sont surtout les constructions *ser* + adjectif + infinitif qui forment des constructions impersonnelles. D'autres moyens d'expression d'impersonnalité peuvent être tirés du sous-corpus : *para* ou *a fin de* + infinitif, les verbes *haber*, *hacer* ou *parecer* ou la forme du gerundio.

Cependant, dans l'échantillon n'apparaît aucune phrase dans laquelle l'équivalent espagnol et tchèque rentrent dans la catégorie des constructions impersonnelles en même temps. La variante tchèque de l'exemple 11 est la suivante : *Když stojíme před příliš velkou záhadou, **neodvážíme se** neuposlechnout.* La première personne du pluriel remplace

le pronom *on* et vu le contexte plus large, c'est une vérité générale. Ainsi, l'espagnol et le tchèque ont recours à un degré de détermination différent. La tendance générale est toutefois opposée : ce sont surtout les occurrences du passif réfléchi qui sont observables au lieu des constructions impersonnelles, le même degré de détermination est d'habitude garanti.

(12) *Il y avait un coin d'où on ne voyait pas la route quand on était assis.* (Boris Vian, *J'irai cracher sur vos tombes*)

- *Bylo tam zákoutí, odkud nebylo od země vidět na silnici.* (traduit par Jaroslav Zaorálek)
- *Había un rincón desde el que no se veía la carretera estando sentado.* (traduit par Jordi Martí Garcés)

Dans l'exemple 12, deux occurrences du pronom *on* trouvent quatre types d'équivalents en tchèque et en espagnol. Dans la phrase tchèque, le premier équivalent est une construction impersonnelle par excellence (*je vidět*), similaire à celle dans l'exemple 10. Le deuxième équivalent demeure plus compliqué à classer : il s'agit d'un groupe prépositionnel en fonction de complément circonstanciel de lieu. Il pourrait être considéré comme un exemple de nominalisation, de construction impersonnelle, voire d'équivalent zéro. Il en résulte que le classement n'est pas du tout évident.

Le premier équivalent rentre sans doute dans la catégorie du passif réfléchi. L'interprétation du deuxième équivalent espagnol mérite une attention particulière car le gerundio *y* est employé. Le gerundio qui complète une proposition principale dont le sujet est explicité, par exemple *Trabajando mucho, puedes llegar lejos*<sup>28</sup>, garde sa valeur référentielle. En revanche, le gerundio tel que celui de l'exemple 12 efface le référent, il peut donc être classé parmi les moyens d'indétermination.

En somme, il en résulte que dans les textes littéraires, c'est surtout le contexte qui permet de placer un équivalent sur l'échelle de la détermination. Les tendances générales qui en ont été dégagées seront présentées dans le chapitre 3.4 et comparées avec celles tirées des textes juridiques. C'est pour cela qu'une analyse basée sur le sous-corpus Acquis et suivant la même structure sera effectuée.

---

<sup>28</sup> De plus, cette phrase peut être comprise de deux manières : (1) comme une explicitation du référent contenu dans le contexte précédent (le degré de détermination le plus élevé) ; ou bien (2) comme une vérité générale si le contexte ne le spécifie pas.

### 3.3 Analyse des équivalents dans les textes juridiques

L'analyse des équivalents du pronom *on* dans les textes juridiques sera effectuée. La procédure suivra le modèle esquissé précédemment. D'abord, la composition des équivalents regroupés en 12 catégories sera présentée. Puis, les correspondances entre les équivalents seront examinées de point de vue quantitatif, mais aussi qualitatif. Finalement, ces équivalents seront placés sur l'échelle de la détermination et comparés du point de vue du degré de détermination.

	Tchèque			Espagnol		
	Rang	Fq absolue	% dans l'échantillon	Rang	Fq absolue	% dans l'échantillon
<b>Passif réfléchi</b>	1	491	46,78 %	1	777	74 %
<b>Construction impersonnelle (<i>lze</i> etc.)</b>	2	146	13,89 %	2	119	11,33 %
<b>Passif promotionnel</b>	3	119	11,33 %	8	12	1,11 %
<b>Nominalisation</b>	4	78	7,44 %	3	41	3,89 %
<b>Équivalent zéro</b>	5	76	7,22 %	3	41	3,89 %
<b>Adjectivisation</b>	7	40	3,78 %	6	18	1,67 %
<b>1<sup>ère</sup> personne du pluriel</b>	8	34	3,22 %	9	8	0,78 %
<b>3<sup>e</sup> personne du pluriel</b>	9	24	2,33 %	5	22	2,11 %
<b><i>Člověk/lidé/gente</i></b>	10	0	0,00 %	10	0	0,00 %
<b><i>Někdo/nikdo/alguno/uno/nadie</i></b>	10	0	0,00 %	10	0	0,00 %
<b>2<sup>e</sup> personne du pluriel/singulier</b>	10	0	0,00 %	10	0	0,00 %
<b>Autres</b>	6	42	4,00 %	7	13	1,22 %

**Tableau 8 :** Composition des équivalents du pronom *on* en tchèque et en espagnol (dans les textes juridiques)

Ce tableau présente douze catégories d'équivalents observées dans le sous-corpus des textes juridiques, classées selon le taux de leur fréquence dans l'échantillon de 1050 occurrences (11 % de la totalité du sous-corpus).<sup>29</sup> La première colonne indique le nom de la catégorie, la deuxième marque le rang de la fréquence de cet équivalent en tchèque, la troisième montre

<sup>29</sup> En comparaison avec le chapitre 3.2, l'ajout des catégories de l'adjectivisation et de nominalisation a été effectué afin de mieux saisir les correspondances, voir le chapitre 1.3 pour plus de détails.

la fréquence absolue des occurrences dans l'échantillon, la quatrième indique le taux de fréquence en tchèque. De même, la cinquième, la sixième et la septième colonne représentent le rang de la fréquence, la fréquence absolue et le taux de fréquence de chaque équivalent en espagnol.

Comme dans le chapitre précédent, les équivalents dont le taux de fréquence équivaut ( $\pm 2\%$ ) sont marqués en jaune. Observons qu'ici, il s'agit de la troisième personne du pluriel, ce qui est totalement opposé à l'observation des textes littéraires dans lesquels c'était surtout la première personne du pluriel, les constructions impersonnelles, les équivalents zéro et les pronoms indéfinis. Ceci sera expliqué par d'une façon plus détaillée dans les analyses des exemples, il faut toutefois souligner dès maintenant que ces types d'équivalents font preuve d'une forte tendance d'impersonnalisation dans les textes juridiques. Ainsi, il est possible de supposer qu'ils ont le plus souvent recours au degré le plus bas de l'échelle de la détermination.

En ce qui concerne les pronoms indéfinis (*člověk* etc., *někdo* etc.), ils sont aussi marqués en jaune mais n'apparaissent jamais dans le sous-corpus Acquis, de même que la deuxième personne du singulier ou pluriel. Aussi, les positions de chaque catégorie sont très distinctes – c'est le passif réfléchi qui prédomine dans les deux langues (en espagnol, il constitue trois quarts des équivalents, en tchèque un peu moins d'une moitié). En revanche, la première personne du pluriel n'est presque jamais utilisée.

Comme dans le chapitre précédent, les équivalents considérablement plus présents dans une langue que dans l'autre sont marqués en rouge, le contraire, c'est-à-dire les équivalents sensiblement moins représentés, en vert. Une tendance d'uniformité des équivalents est visible en espagnol<sup>30</sup>, seuls 26 % des occurrences ne sont pas traduits par la troisième personne du pluriel. En revanche, les équivalents tchèques semblent plus proportionnellement étendus, dont fait preuve par exemple le fait que les catégories à la troisième position en espagnol (nominalisation et équivalent zéro) sont presque aussi fréquentes que la catégorie à la septième position en tchèque (adjectivisation). De plus, en comparaison avec les textes littéraires, les équivalents du pronom *on* dans les textes juridiques sont moins variés.

---

<sup>30</sup> Dans ce sens, le terme « *uniformité* » est employé pour dénoter une forte prédominance d'un équivalent.



Afin de mieux révéler si les procédés de traduction sont réellement si différents que le tableau suggère, il est approprié de voir si les catégories se correspondent en espagnol et en tchèque.

### 3.3.1 Correspondances entre les équivalents du pronom *on* en tchèque et en espagnol dans les textes juridiques

Après avoir analysé et comparé la composition des équivalents en tchèque et en espagnol, il faut maintenant prêter attention à la correspondance entre ces équivalents. C'est pour cela que deux tableaux vont être constitués afin de démontrer dans quelles catégories les langues se correspondent et dans quelles catégories elles ne trouvent aucune intersection.

	Moyenne dans l'échantillon	Proportion de la totalité
<b>Passif réfléchi</b>	61	40,67 %
<b>Construction impersonnelle (<i>lze</i> etc.)</b>	6,14	4,10 %
<b>Nominalisation</b>	2,71	1,81 %
<b>Équivalent zéro</b>	2,57	1,71 %
<b>Adjectivisation</b>	1,43	0,95 %
<b>Passif promotionnel</b>	0,43	0,29 %
<b>3<sup>e</sup> personne du pluriel</b>	0,43	0,29 %
<b>1<sup>ère</sup> personne du pluriel</b>	0,29	0,19 %
<b>2<sup>e</sup> personne du pluriel/singulier</b>	0	0,0 %
<b>Pronom indéfini</b>	0	0,0 %
<b>Correspondances au total</b>	75,14	50,1 %

**Tableau 9 :** Correspondance entre les équivalents – proportion de la totalité de l'échantillon (dans les textes juridiques)

Voici le tableau semblable à celui du chapitre précédent. Il présente le nombre moyen de correspondances entre les équivalents tchèques et espagnols dans sept échantillons des textes juridiques dont l'étendue est 150 occurrences. La première colonne dénote le type d'équivalent. Dans la deuxième colonne, la moyenne de nombre de correspondances dans l'échantillon est marquée. La troisième colonne indique la proportion des correspondances dans la totalité de l'échantillon.

Remarquons que le nombre total des correspondances est légèrement supérieur à celui des textes littéraires, 50,1 %. Il faut toutefois prendre en considération que plus de 40 % sont des correspondances du passif réfléchi, ce qui témoigne de la tendance d'uniformité (c'est-à-dire la prédominance de l'équivalent appartenant à la catégorie du passif réfléchi) mentionnée ci-dessus. Contrairement aux textes littéraires, la correspondance entre les équivalents dans la catégorie des constructions impersonnelles est assez fréquente (n'oublions pas que dans le sous-corpus Core, aucune correspondance dans cette catégorie n'a été trouvée), elles sont même placées à la deuxième position selon la fréquence. Conformément aux attentes, la deuxième personne du singulier ou du pluriel n'est pas employée vu qu'il s'agit des textes applicables sur tous les citoyens de l'Union européenne, pas sur des personnes concrètes, ou bien sur les créateurs des textes.<sup>31</sup> Ces observations peuvent être mises en relation avec le chapitre 3.1.3 où la prédominance du verbe *entendre* en voisinage du pronom *on* a été remarquée. Ainsi, il est évident que la locution « *on entend* » ne se référera à un référent précis, voire l'interlocuteur.

	Nº d'occurrences (réunion tchèque + espagnol)	Nº de correspondances (intersection)	Quotient
<b>Passif réfléchi</b>	769	427	56 %
<b>Construction impersonnelle (<i>Ize</i> etc.)</b>	207	43	21 %
<b>Nominalisation</b>	94	19	20 %
<b>Équivalent zéro</b>	92	18	20 %
<b>Adjectivisation</b>	44	10	23 %
<b>Passif promotionnel</b>	120	3	3 %
<b>3<sup>e</sup> personne du pluriel</b>	41	3	7 %
<b>1<sup>ère</sup> personne du pluriel</b>	32	2	6 %
<b>2<sup>e</sup> personne du pluriel/singulier</b>	0	0	0 %
<b>Pronom indéfini</b>	0	0	0 %

**Tableau 10 :** Correspondance des catégories d'équivalents en tchèque et en espagnol (dans les textes juridiques)

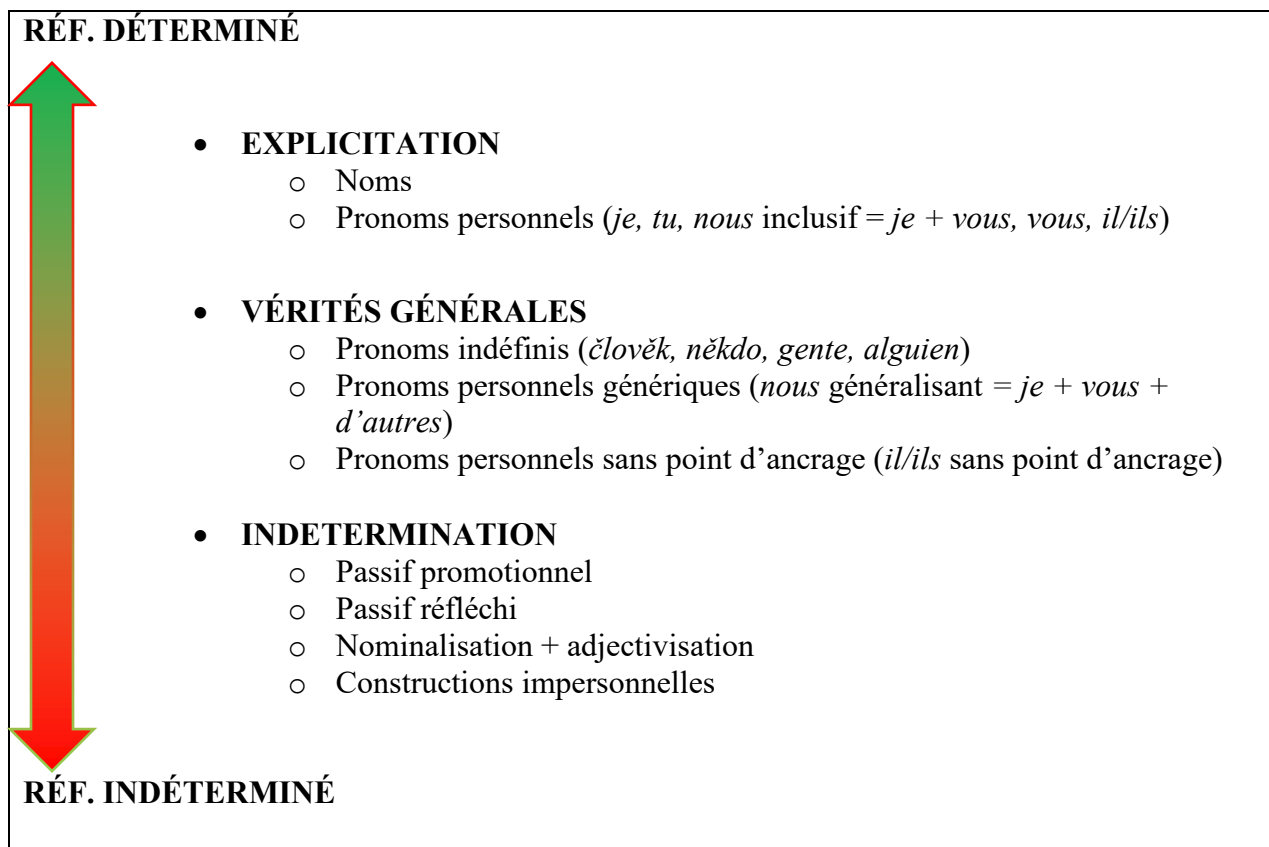
<sup>31</sup> Cette idée va être plus développée dans le chapitre 3.3.2 qui traite la problématique de l'échelle de la détermination.

Comme le tableau 8, le tableau 10 montre d'abord le nombre d'occurrences de chaque catégorie en tchèque et en espagnol, puis le nombre d'occurrences dans lesquelles la catégorie d'équivalent dans les deux langues se correspondent. La quatrième colonne est le résultat de la division (en forme de pourcentage arrondi) où le numérateur est la troisième colonne (c'est-à-dire l'intersection des correspondances des équivalents tchèques et espagnols) et le dénominateur est la deuxième colonne (c'est-à-dire la réunion desdits équivalents).

Il est évident que c'est le passif réfléchi où le tchèque et l'espagnol se correspondent le plus. Cependant, dans les textes littéraires, il y a deux catégories avec un quotient semblable, ce qui fait voir encore une fois la tendance de l'uniformité dans les textes juridiques. En revanche, quatre catégories suivantes montrent un degré de correspondance deux fois plus élevé que les catégories dans les textes littéraires. La grande différence entre la réunion et l'intersection dans le cas du passif promotionnel ne devrait pas rester inaperçue non plus. Ce procédé est assez fréquent, les deux langues n'y ont pourtant pas recours dans les mêmes phrases. Afin d'observer la tendance encore plus générale à déterminer le référent, il faut maintenant placer les exemples sur l'échelle de la détermination car c'est surtout le degré de détermination qui est intéressant du point de vue contrastif.

### **3.3.2 Echelle de la détermination dans les textes juridiques**

Les équivalents peuvent être regroupés en 3 catégories sur l'échelle de la détermination. En vert sont marqués les types les plus explicites. Le passage de transition entre le vert et le rouge signifie les équivalents qui dénotent des vérités générales applicables sur tous (voir le chapitre 1.2.1), dont le référent peut être aussi bien déterminé qu'indéterminé. En bas de l'échelle, en rouge, sont indiqués les moyens d'expression de l'indétermination totale.



**Figure 4 : L'échelle de la détermination (dans les textes juridiques)**

Pour obtenir des résultats fiables, il faut légèrement modifier l'échelle de la détermination introduite dans le chapitre précédent car deux procédés fréquents s'introduisent dans le sous-corpus Acquis.<sup>32</sup> Il s'agit de la nominalisation et de l'adjectivisation.<sup>33</sup> Ces procédés sont classés parmi les types dont la détermination est moins évidente vu qu'ils n'explicitent pas le référent et font en sorte que son explicitation ne soit pas nécessaire. Par exemple, il est impossible de distinguer le référent du pronom *on* dans la locution : « *des évolutions des flux de conteneurs que l'on a décrites précédemment* », traduite en tchèque comme « *nejen vzhledem k výše popsaným trendům v oblasti kontejnerové dopravy* » et en espagnol comme « *la evolución del transporte por contenedores arriba descrita* ». Cependant, examinons d'abord les exemples de l'explicitation.

<sup>32</sup> Bien sûr, ces deux procédés ont apparu déjà dans le sous-corpus des textes littéraires. Leurs occurrences ont cependant été si peu fréquentes qu'ils sont classés dans la catégorie *autres* afin de ne pas surcharger le tableau.

<sup>33</sup> Parmi les exemples de l'adjectivisation sont classées aussi les formes du participe passé adjectival en fonction d'épithète.

### 3.3.2.1 Explicitation

L'explicitation n'est pas un procédé typique des équivalents tchèques et espagnols dans le contexte des textes juridiques. Pourtant, des exemples de ce degré peuvent être trouvés même dans le sous-corpus Acquis.

#### I) Noms

En premier lieu, il faut évoquer que contrairement aux textes littéraires, le sous-corpus Acquis ne contient pas beaucoup d'exemples de référents explicités par l'emploi des noms. Cependant, dans les cas observés, le pronom *on* est parfois explicité par le nom de l'institution concernée par le document législatif ou de l'institution qui l'a créé, dont fait preuve l'exemple suivant :

(13) **On rappelle** à ce sujet l'avis exprimé par le Comité des régions sur la communication de la Commission sur "Accroître l'emploi des travailleurs âgés et différer la sortie du marché du travail" (adopté au cours de la 56ème session plénière, qui s'est tenue le 29 septembre 2004).

- *Výbor by chtěl odkázat na své stanovisko (přijaté v průběhu plenárního zasedání dne 29. září 2004) ke sdělení Komise týkající se Zvyšování zaměstnanosti starších pracovníků a prodlužování odchodu z trhu práce.*

L'exemple 13 est un exemple prototypique du procédé de l'explicitation directe du référent. Vu qu'il s'agit d'un document intitulé Avis du Comité des régions, il est sans doute irréprochable de supposer que le référent de *on* est ce comité, ce que le tchèque transpose en explicitant ce référent, contrairement à l'espagnol qui a recours à une construction impersonnelle. Il est possible de déduire que l'espagnol insiste sur la valeur générale de l'énoncé contrairement au tchèque qui préfère l'explicitation de l'énonciateur. Néanmoins, il ne s'agit pas d'une tendance observable dans tous les cas, dont fait preuve l'exemple 14.

(14) Territorialité : *on transcende le concept de proximité territoriale et on le remplace par un concept de relation de réseau et de valorisation.*

- *Lokalita: koncepci zeměpisné blízkosti nahrazuje koncepce síťových vazeb a maximálního využívání různých kulturních identit partnerů.*

L'exemple 14 semble plus éloigné de l'explicitation directe du référent car le référent du pronom *on* est en réalité effacé. En effet, le sujet de la phrase tchèque, *koncepce*, n'est

pas un actant qui effectue l'action (un agent), son rôle est plutôt celui d'un patient. Or, l'ordre des mots en phrase tchèque étant variable, il offre une palette plus vaste d'équivalents et un pré carré pour la modulation. Dans ce cas, il ne s'agit pas d'un glissement de sens à proprement parler mais plutôt d'une interprétation qui va un peu plus loin que l'original, pourtant pas trop loin. Elle garde toujours le sens de la phrase initiale.

Néanmoins, en essayant d'attribuer un degré de détermination à cet exemple, il faut plutôt chercher parmi les degrés les plus bas car le référent du pronom *on* n'est pas du tout explicité. Au contraire, il semble plus proche des constructions impersonnelles bien que cet exemple fasse théoriquement partie de la catégorie des noms classée parmi les équivalents de l'explicitation. Il est donc compliqué de placer cet exemple sur l'échelle de la détermination, surtout dans les cas des exemples déviants comme l'exemple 14.

## II) Pronoms personnels (je, tu, nous inclusif = je + vous, vous, il/ils)

Les pronoms personnels sont une catégorie qui est très difficile à placer sur l'échelle de la détermination vu qu'ils peuvent apparaître sans ou avec un point d'ancrage. De nouveau, rappelons que c'est le contexte qui précise s'il s'agit du premier ou du second cas. Les pronoms personnels sont ce second cas mais n'apparaissent que peu dans le sous-corpus Acquis.

(15) *Dans le cadre des stratégies de développement rural appliquées à la méthode LEADER, l'on agit avec une grande efficacité sur le patrimoine rural, compris dans son sens le plus large.*<sup>34</sup>

- *Ženy hrají významnou úlohu v diversifikaci venkovského hospodářství, často ale čelí specifickým překážkám, jako je omezený přístup ke kapitálu. Pozměňovací návrh je zaměřený tak, aby podpořil opatření k odstranění těchto překážek. V rámci strategií rozvoje venkova EAFRD úspěšně **podporují** venkovské dědictví v širším slova smyslu.*

Afin d'éclairer le choix de traduction et de détermination en tchèque, un contexte plus large doit être ajouté. C'est pour cela qu'au lieu d'une phrase, trois phrases sont citées dans la version tchèque. Ceci fait preuve du fait déjà évoqué : il est essentiel de prendre en considération le contexte large. Ainsi, le tchèque explicite le référent *ženy* sans que

<sup>34</sup> Remarquons aussi que l'expression « *l'on agit* » insère le « *l'* » pour des raisons phonétiques introduites dans le chapitre 1.1.2, surtout l'insertion après la pause.

l'interprétation soit erronée. Néanmoins, cette explicitation reste très peu fréquente, aucune occurrence de la première et de la deuxième personne du singulier n'est observée. En ce qui concerne le pronom personnel *nous*, il est le plus souvent employé comme un procédé de la vérité générale.

### 3.3.2.2 Vérités générales

Le second degré de l'échelle de la détermination est représenté par les vérités générales. Comme indiqué précédemment, les pronoms indéfinis (tels que *člověk*, *la gente*, *někdo* ou *alguien*) n'apparaissent pas dans les textes juridiques. Il serait donc inutile d'introduire cette catégorie dans l'analyse. Cette partie se penche donc sur l'étude des pronoms personnels génériques et sans point d'ancrage.

#### I) Pronoms personnels génériques (nous généralisant = je + vous + d'autres)

Contrairement aux textes littéraires, les textes juridiques n'emploient pas souvent la première personne du pluriel, ce qui n'est pas si étonnant vu que l'objectif de ces textes n'est pas de s'identifier avec le lecteur ou la population en général. Néanmoins, le tchèque a parfois recours au pronom personnel *nous* généralisant qui est complémentaire à la catégorie du pronom personnel *nous* inclusif. Ce type n'apparaît pourtant pas dans les textes juridiques.

(16) *L'on peut voir que cette approche de la Commission [...].*

- *Budeme si moci všimnout, že toto ustanovení Komise [...].*
- *Más adelante observaremos que esta postura de la Comisión [...].*

Dans l'exemple 16, prêtons attention aussi au temps verbal des équivalents : le futur. La valeur du futur en espagnol est mieux développée dans l'exemple 19. Il faut toutefois remarquer que dans ce cas, le futur en tchèque n'a pas la même valeur – celle de réduire la certitude de l'accomplissement de l'action. En ce qui concerne l'interprétation de cet emploi, il est probable qu'il s'agit d'une prolepse d'une observation qui suivra plus tard et que la version française voit comme réelle alors que les traductions tchèque et espagnole y rajoutent un peu d'incertitude.

## II) Pronoms personnels sans point d’ancrage (*il/ils* sans point d’ancrage) et pronoms indéfinis (*člověk, někdo, la gente, alguien* etc.)

Contrairement aux attentes, les pronoms personnels de la troisième personne du pluriel ou singulier sans point d’ancrage n’apparaissent jamais dans les textes juridiques alors qu’ils sont utilisés assez fréquemment dans les textes littéraires. De plus, le langage dans le sous-corpus Acquis ne contient aucune occurrence des pronoms indéfinis tels que *člověk*, *la gente* ou *alguien*. En revanche, ce type d’équivalent est très fréquent (surtout en tchèque) dans les textes littéraires.

### 3.3.2.3 Indétermination

Comme prévu, l’indétermination est le degré le plus répandu parmi les référents du pronom *on* dans les textes juridiques. Cette observation n’est pas surprenante vu que l’objectif des textes législatifs n’est pas d’identifier un référent concret. Voici donc les différentes possibilités de l’expression du degré le plus bas de détermination.

#### I) Passif promotionnel

Le passif promotionnel est un procédé souvent utilisé en tchèque (dont témoigne la première partie de ce chapitre, voir le sous-chapitre 3.3.1). L’exemple 17 est un exemple prototypique de l’usage du passif promotionnel en tchèque. Le plus souvent, c’est le passif réfléchi qui lui correspond en espagnol.

(17) *Le producteur ou le fabricant d'un produit alimentaire doit décider si l'on peut consommer ce produit tel quel [...].*

- *Producent nebo výrobce potravinářského produktu musí rozhodnout, zda může být produkt konzumován jako takový [...].*
- *El productor o el fabricante de un producto alimenticio debe decidir si el producto está listo para ser consumido como tal [...].*

Cependant, dans quelques cas, l’équivalent espagnol lui correspond aussi, comme dans le cas de l’exemple 17. Le passif promotionnel est une illustration par excellence de l’indétermination du référent, il est totalement omis. Il peut théoriquement être complété par un actant ou agent, ce qui ne se passe toutefois presque jamais dans les textes juridiques étant que le référent cesserait d’être vague et indéterminé.



## II) Passif réfléchi

Le chapitre sur les correspondances entre les équivalents du pronom *on* en tchèque et en espagnol a montré que la plus grande quantité des correspondances peut être observée parmi les occurrences du passif réfléchi. Il s'agit du procédé le plus souvent employé pour saisir le degré de détermination le moins élevé.

(18) *À cette fin, on entend par « impuretés végétales » des matières végétales visibles [...].*

- *Pro účely tohoto odstavce se "rostlinnými nečistotami" rozumějí části rostlin rozeznatelné pouhým okem [...].*
- *A estos efectos, se entiende por « impurezas de origen vegetal » materias de origen vegetal perceptibles a simple vista [...].*

La fréquence élevée du passif réfléchi (46 % en tchèque et 74 % en espagnol) est surtout causée par la fréquence élevée du verbe *entendre* dans le sous-corpus Acquis. L'uniformité est parfois déséquilibrée par l'équivalent de la classe du passif promotionnel ou de la première personne du pluriel. De plus, il est intéressant de mentionner que la formule « *on entend par* » se traduit en espagnol parfois comme « *se entiende* » au présent (exemple 18), parfois comme « *se entenderá* » au futur (exemple 19).

(18) *À cette fin, on entend par « impuretés végétales »*

- *A estos efectos, se entiende por « impurezas de origen vegetal » [...].*

(19) *Aux fins du présent règlement, on entend par « carcasse de porc »*

- *A los fines del presente Reglamento, se entenderá por « canal de cerdo » [...].*

Cependant, la fréquence de la deuxième possibilité (la forme du futur, exemple 19) est supérieure à celle de la première (la forme du présent, exemple 18). Dans la totalité du sous-corpus Acquis aligné français-tchèque-espagnol, « *on entend* » est traduit 2861 fois par la forme du futur, alors que le présent n'apparaît que 383 fois.

Ceci peut être expliqué en examinant la forme du futur en espagnol du point de vue onomasiologique. Ce temps verbal peut désigner à la fois un événement futur dont l'accomplissement n'est pas complètement garanti, à la fois le présent du probabilif ; le futur donc exprime un degré moins élevé de certitude que le présent de l'indicatif.<sup>35</sup> Il en résulte que dans ce cas, le futur est doté d'une capacité de montrer que l'action est fort

---

<sup>35</sup> Le probabilif est le terme employé par Čermák et Zavadil (2005, p. 290–305) pour désigner une forme d'expression de la signification modale de la probabilité.

probable, mais il se peut qu'elle change dans l'avenir. En tout cas, le futur est typique pour les définitions préliminaires des termes introduits dans les règlements. De plus, la fréquence élevée du futur dans les textes juridiques espagnols peut-être aussi provoquée par son double-jeu de modalité.<sup>36</sup>

### III) Nominalisation et adjectivisation

La nominalisation et l'adjectivisation sont deux procédés qui fonctionnent aussi dans les textes littéraires, leur fréquence y est toutefois presque négligeable. En revanche, les textes juridiques y ont recours assez souvent, dont témoigne le tableau 10.

(20) *Si l'on voulait poursuivre la logique de l'argumentation de l'Allemagne, les créanciers privés devraient davantage renoncer à leurs créances que les créanciers publics parce que ces abandons de créances sont déductibles de leur base imposable.*<sup>37</sup>

- *Z logického rozvíjení argumentace Německa vyplývá, že věřitelé ze soukromého sektoru by museli prominout vyšší dlužnou částku než věřitelé z veřejného sektoru, protože jejich prominutí není zahrnuto do daňového výměru.*

(21) *Toutefois, compte tenu des informations dont on dispose, il convient de limiter l'exposition à l'oestradiol [...].*

- *Vzhledem ke stávajícím informacím je zapotřebí omezit expozici organismu účinkům 17-beta-estradiolu [...].*
- *Sin embargo, a la luz de la información existente es conveniente limitar la exposición al estradiol-17  $\beta$  [...].*

Le tchèque et partiellement l'espagnol emploient la nominalisation et l'adjectivisation surtout dans le cas des verbes suivants : *comparer* (ve srovnání, en comparación), *disposer* (být k dispozici, stávající, existente), *évaluer* (posuzování, hodnocení, evaluación), *risquer* (nebezpečí, riesgo), *appeler* (takzvaný) ou *exclure* (vylučující, denominado). La nominalisation et l'adjectivisation permettent de ne pas préciser le référent du pronom *on* tout en gardant le sens du verbe, c'est pour cela que ces deux procédés sont assez souvent

<sup>36</sup> Il est remarquable que ce sont les textes juridiques où apparaît aussi le futur du subjonctif, presque disparu de l'espagnol moderne.

<sup>37</sup> Dans l'exemple 20, notons que l'équivalent tchèque correspond au prédicat complexe *vouloir poursuivre*, non seulement au verbe modal *vouloir*.

employés dans les deux langues. De plus, l'exemple 20 témoigne d'un dépouillement qui change la modalité de la phrase originale – la phrase conditionnelle introduite par *si* ainsi que le verbe *vouloir* sont effacés dans la traduction tchèque.

#### IV) Constructions impersonnelles

La construction impersonnelle est placée à la deuxième position selon la fréquence dans les deux langues dans les textes juridiques. En tchèque, ce sont surtout les expressions telles que *lze*, *je třeba*, *je možné*, *je nutné* qui forment la plupart des équivalents. Cependant, en général, la forme de ces équivalents est la suivante : le verbe *být* + adjectif.

(22) *L'on devrait tenir compte du fait que dans les nouveaux pays membres d'Europe centrale et orientale, l'infrastructure est en mauvais état [...].*

- *Je třeba mít na paměti, že v nových členských státech ve střední a východní Evropě je infrastruktura v horším stavu [...].*
- *Hay que tener en cuenta asimismo que las infraestructuras en los nuevos Estados miembros de Europa Central y Oriental están en peor estado [...].*

Bien sûr, l'espagnol aussi peut créer des constructions impersonnelles de cette forme. Néanmoins, contrairement au tchèque, l'espagnol dispose de moyens d'impersonnalisation plus variés. Par exemple, il existe des verbes dont la forme reste impersonnelle – *cabere*, *haber* et *quedar*. Ce type de traduction est souvent lié aux verbes comme *espérer*, *prévoir* ou *mentionner*. Un autre procédé important des équivalents espagnols impersonnels est la construction *para/a fin de* + infinitif<sup>38</sup> (ce qui est souvent la traduction de la locution « *on veut* »), l'infinitif en soi représente aussi un moyen comment rendre le référent invisible. Finalement, c'est le gerundio<sup>39</sup> (dont les occurrences ne sont pas nombreuses, pourtant pas négligeables) qui n'explicite pas le référent.

---

<sup>38</sup> A titre d'exemple, mentionnons l'équivalent *a fin de* de la proposition conditionnelle suivante : « *Si l'on veut assurer la libre circulation de tous ces médias [...]* » « *A fin de garantizar la libre circulación de estos medios de comunicación [...]* ». Il est à constater que dans ce cas, le verbe *vouloir* est effacé, la modalité est donc modifiée et le sens légèrement glisse.

<sup>39</sup> C'est aussi le gerundio qui est souvent employé en tant qu'équivalent espagnol des propositions conditionnelles. Voici un exemple de son usage : « *Si l'on retient 2002 comme année de référence [...]* » « *Utilizando 2002 como año de referencia [...]* ». Il est à noter que la valeur modale et la proposition conditionnelle de l'original français ne se perdent car le gerundio garde ces qualités.

### 3.3.2.4 Équivalent zéro

Finalement, une dernière catégorie doit être introduite. Il est impossible de le placer dans le groupe de la détermination car le référent n'est pas explicité, il est omis. De l'autre côté, il ne correspond pas au groupe de l'indétermination non plus car il ne s'agit pas d'un moyen d'indétermination à proprement parler. C'est pour cela que cette catégorie ne sera pas prise en compte dans la composition de l'échelle de la détermination (l'addition des pourcentages n'équivaudra donc pas à 100 %). Cependant, comme l'équivalent zéro est la cinquième catégorie la plus représentée en tchèque et la troisième en espagnol dans les textes juridiques, ne pas la mentionner serait incorrect.

*(23) En ce qui concerne le volume des ventes des autres producteurs communautaires, on observe une diminution de 7,1 % en volume et de 7,8 % en parts de marché entre 2001 et la période d'enquête.*

- *Pokud jde o objemy prodeje ostatních výrobců ve Společenství, Ø poklesly mezi rokem 2001 a obdobím šetření o 7,1 % a jejich podíl na trhu se ve stejném období snížil o 7,8 %.*
- *Por lo que respecta a los volúmenes de venta de otros productores comunitarios, Ø estos han descendido un 7,1 % en volumen entre 2001 y el período de investigación y un 7,8 % en cuota de mercado durante el mismo período.*

L'exemple 23 montre bien que l'omission de traduction n'engendre pas un glissement de sens. C'est ici qu'il faut rappeler le chapitre 3.1 où les fréquences des verbes qui accompagnent le pronom *on* ont été étudiées.

L'omission totale en tchèque est fréquente dans le cas du verbe *observer*.<sup>40</sup> Seuls 20 % des occurrences dans la totalité du sous-corpus Acquis seraient classées dans la catégorie des constructions impersonnelles (par exemple *lze* ou *je možné*), la première personne du pluriel n'apparaît que quatre fois, l'adjectivisation seulement deux fois. Il en résulte que 60 % des équivalents tchèques sont des équivalents zéro. De l'autre côté, dans ce cas, l'espagnol a recours à utiliser le passif réfléchi (dans 67 % des occurrences),

---

<sup>40</sup> C'est pour cela que le chapitre sur l'accompagnement lexical du pronom *on* (le chapitre 3.1.3) a été incorporé dans ce travail. Selon les données, il est évident que dans certains cas, un verbe (bien qu'il s'agisse d'un verbe à sens plein) peut être omis dans la traduction sans que la phrase perde son sens initial.

la première personne du pluriel ainsi que les constructions impersonnelles ne sont utilisées qu'une fois et le reste, c'est-à-dire 28 %, sont des équivalents zéro.

Outre le verbe *observer*, c'est aussi le verbe *retenir* qui influence la composition des traductions en tchèque car son équivalent est souvent zéro. Cependant, cela ne veut pas dire que le nombre assez élevé des occurrences de l'équivalent zéro signifie une tendance générale d'omettre la traduction. Bien au contraire, dans l'échantillon analysé, les textes juridiques ne manquent presque jamais d'équivalent. A la même tendance obéissent les verbes comme *appeler*, *mentionner* ou *citer*, c'est aussi la locution « *on peut dire/penser* » dont les équivalents tchèques et espagnols sont souvent zéro.<sup>41</sup> Cette observation est importante : il semble que le référent du pronom *on*, ou plutôt le degré de sa détermination, n'est pas toujours si important, surtout en tchèque.

De nombreux exemples ont été observés au cours des deux chapitres précédents, chacun situé dans un des trois degrés de la détermination : (1) l'explicitation, (2) les vérités générales, (3) l'indétermination. Maintenant, afin de dégager les tendances générales du fonctionnement du pronom *on* dans les textes littéraires et juridiques, une comparaison des échelles de la détermination sera présentée.

---

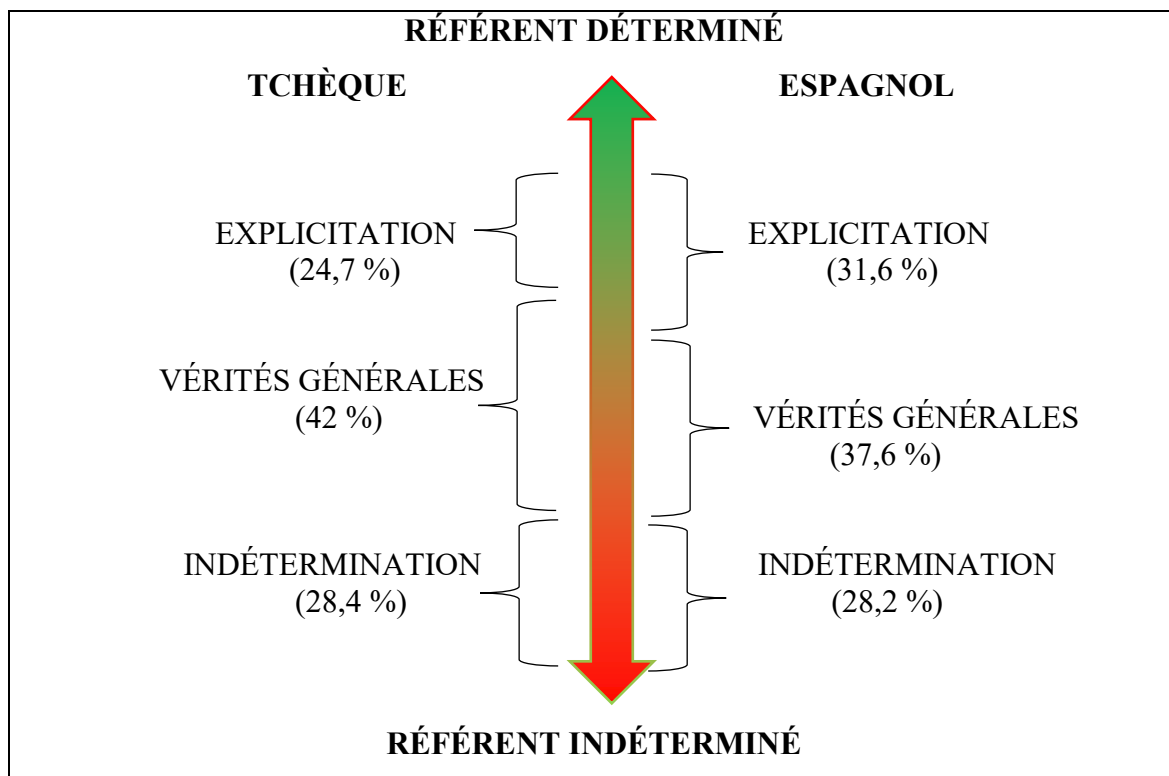
<sup>41</sup> De cette remarque fait preuve par exemple la phrase française et ses équivalents suivants : « *En ce sens, on peut dire que le Land avait déjà adopté [...].* » « *V této míře Ø se země již chovala [...].* » « *A este respecto, Ø el Land actuó ya como un inversor privado [...].* »

### 3.4 Comparaison des échelles de la détermination

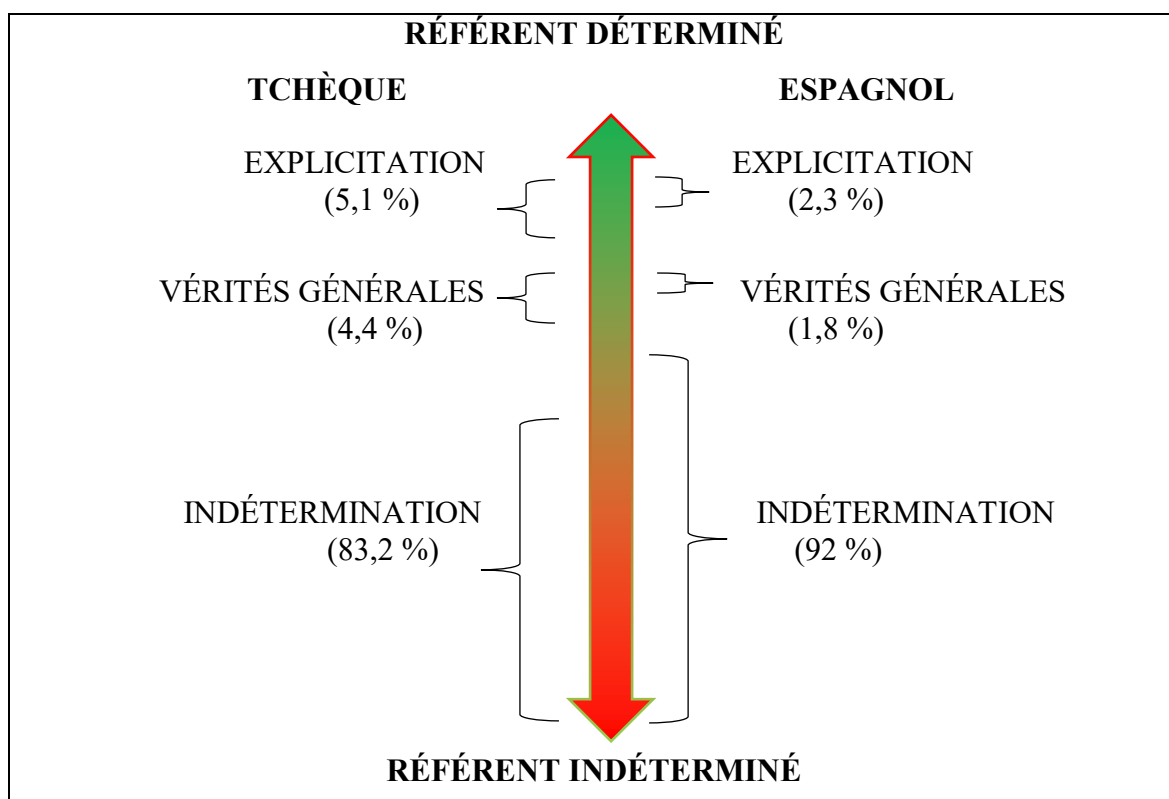
Le chapitre précédent a montré les correspondances entre chaque catégorie des équivalents du pronom *on* en tchèque et en espagnol ainsi que les différents exemples de l'échelle de la détermination. Néanmoins, en ce lieu, il faut s'intéresser non seulement aux correspondances entre les catégories, mais surtout entre les degrés de détermination.

Après avoir regroupé les équivalents, il semble que le taux des correspondances entre les degrés de détermination est plus élevé que le taux de correspondances entre les différentes catégories d'équivalents. Pour ce qui est du degré de détermination dans les textes littéraires, 54,3 % des occurrences du pronom *on* se correspondent en tchèque et en espagnol. Dans les textes juridiques, ce pourcentage s'élève à 60,8 % des occurrences.

Ces chiffres indiquent que bien que le degré de détermination ne soit pas objectivement fixé, les correspondances demeurent assez nombreuses, ce n'est donc pas que le contexte immédiat qui influence la détermination du référent, mais aussi le contexte plus large et extralinguistique. Cependant, il convient d'introduire une comparaison des échelles de la détermination (figures 5 et 6) qui montre à quel point la composition des degrés de détermination des équivalents est similaire en tchèque et en espagnol.



**Figure 5 :** *Comparaison des échelles de la détermination dans les textes littéraires*



**Figure 6 :** *Comparaison des échelles de la détermination dans les textes juridiques*

Voici deux figures récapitulatives des échelles de la détermination qui comparent la composition des équivalents dans les textes littéraires et juridiques. Les deux figures sont délibérément placées l'une sous l'autre afin de mieux voir les différentes tendances qui distinguent non seulement les deux langues (le tchèque et l'espagnol), mais aussi les deux types de textes (littéraires et juridiques). L'approche scalaire permet de dégager les tendances générales sans prendre le risque d'adopter un classement trop étroit.

La figure numéro 5 fait preuve d'une certaine homogénéité lors de l'emploi de chaque degré de détermination dans les textes littéraires en tchèque et en espagnol. Le terme l'homogénéité signifie que les fréquences des trois degrés sont, dans une certaine mesure, analogues, chacun comportant approximativement un tiers. Dans les deux langues, les vérités générales légèrement prédominent – cette prédominance est un peu plus visible en tchèque (42 %) qu'en espagnol (37,6 %). La fréquence d'utilisation des deux degrés restants, c'est-à-dire l'explicitation et l'indétermination, semble plus ou moins égale.

En revanche, la figure numéro 6 fait voir que les types d'équivalents du pronom *on* dans le sous-corpus Acquis sont plus hétérogènes. Ici, le terme l'hétérogénéité signifie qu'un degré de détermination prédomine considérablement. Ainsi, le référent du pronom *on* dans les textes juridiques est le plus souvent indéterminé – cette catégorie représente 83,2 % des occurrences en tchèque et même 92 % des occurrences en espagnol. C'est pour cela que les types d'équivalents en tchèque semblent être légèrement plus variés.

En ce qui concerne la comparaison des degrés de détermination en tchèque et en espagnol, remarquons qu'au total dans les textes littéraires, les deux langues ont une tendance à employer des équivalents dont le degré est semblable. La fréquence d'équivalents indéterminés ne diffère presque pas, celle des vérités générales est légèrement supérieure en tchèque (de 4,4 %). L'explicitation est le procédé qui est davantage employé en espagnol (il excède le tchèque par 6,9 %). Il est donc évident que les équivalents tchèques peuvent être placés plus bas sur l'échelle alors que les équivalents espagnols ont tendance à être plus explicites.

De l'autre côté, dans les textes juridiques, l'explicitation et la vérité générale sont les degrés où le tchèque dépasse l'espagnol (plus de deux fois). Ils sont toutefois très peu représentés – il a déjà été indiqué que le référent dans le sous-corpus Acquis est presque exclusivement indéterminé vu que l'objectif des textes législatifs n'est pas d'identifier un référent concret.



Cependant, il ne faut pas négliger que ces échelles et la fréquence des trois degrés de détermination ont été élaborées sur les données extraites des échantillons (représentatifs, pourtant pas englobant la totalité du corpus). Il existe donc la possibilité que les chiffres soient modifiés après une étude plus détaillée, basée sur des échantillons plus larges ou suivant une autre méthodologie.

Aussi, classer l'omission de traduction directe dans une des cases est une tâche très compliquée parce que l'équivalent zéro n'est ni déterminé, ni indéterminé. De plus, cette échelle n'est pas utile à cent pour cent car de nombreux exemples sont à la limite entre les catégories. Par exemple, afin d'attribuer une valeur déterminative aux phrases dont l'équivalent du pronom *on* était la première personne du pluriel, il a fallu essayer de déterminer le référent plus concret, c'est-à-dire un *nous* généralisant (*je + vous + d'autres*, qui fait donc partie des vérités générales) ou bien un *nous* concret (*je + vous*, qui est classé parmi les moyens d'explicitation étant donné qu'il peut être déduit à partir du contexte). La troisième personne du pluriel pose le même problème – elle peut être à la fois anaphorique ou déictique (avec un référent explicite) et générale (sans point d'ancrage).

En somme, il en résulte que le degré de détermination du référent du pronom *on* n'est pas complètement indiscutable. Par l'intermédiaire des exemples tchèques et espagnols traduisant les mêmes phrases françaises, il a été démontré qu'il existe toutefois quelques pistes pour l'interpréter. Après avoir classé ces équivalents dans les degrés de la détermination, le taux de correspondances beaucoup plus élevé a été observé.

De plus, notons que c'est aussi le contexte linguistique qui définit, dans la plupart des cas, notre réflexion sur la détermination du référent. Il a déjà été indiqué que ce n'est pas nécessairement le contexte immédiat du pronom *on*, il faut parfois recourir à un passage plus large, voir un chapitre ou même toute une œuvre dans les cas des textes littéraires. Néanmoins, cette tâche semble plus compliquée dans les textes juridiques qui tendent à être moins abondants en contexte, plus économiques, plus stéréotypiques et plus claires (voir le chapitre 1.1.1).

Ainsi, ceci implique que ce n'est pas que le contexte linguistique immédiat ou large qui joue un rôle décisif lors de la détermination des valeurs référentielles du pronom *on*. C'est aussi son contexte extralinguistique, par exemple l'objectif du texte, son type ou le destinataire.

## 4 CONCLUSION

Ce mémoire avait pour but d'analyser les types d'équivalents du pronom *on* dans les traductions en tchèque et en espagnol, en fonction du degré de détermination du référent de ce pronom. D'abord, l'hypothèse sur une fréquence relative du pronom *on* plus élevée dans les textes littéraires que dans les textes juridiques a été confirmée. Dans le sous-corpus Core, la fréquence observée est égale à 804,35 ipm, tandis que la fréquence des occurrences du pronom *on* dans le sous-corpus Acquis est 306,03 ipm. Ainsi, la fréquence du pronom *on* dans les textes littéraires est plus que 2,5 fois supérieure à celle dans les textes juridiques.

La deuxième hypothèse supposant que la diversité lexicale limitée du sous-corpus Acquis influencera la diversité référentielle du pronom *on* a été partiellement confirmée. Effectivement, il a été prouvé que les textes juridiques sont plus uniformes en ce qui concerne le choix d'équivalents vu que c'est l'original qui est plus uniforme en ce qui concerne les types de référents du pronom *on*. Cependant, il n'est pas totalement certain que le vocabulaire limité et la diversité référentielle du pronom *on* limité ne soient pas que deux tendances parallèles dans les textes juridiques (qui ont tendance à être plus objectifs et univoques) plutôt que deux tendances liées par la relation entre une cause et un effet.

Troisièmement, l'étude a confirmé l'hypothèse concernant la fréquence relativement basse de l'accord de l'adjectif ou du participe passé avec le verbe *être*. L'accord ne s'effectue presque jamais dans les textes juridiques, peut-être aussi afin de ne pas expliciter le référent. Dans les textes littéraires, c'est surtout dans *Voyage au bout de la nuit* de Louis-Ferdinand Céline où il est possible de trouver l'accord en genre et/ou en nombre de l'adjectif ou du participe passé. Pourtant, il a été observé qu'aucune des œuvres n'a systématiquement employé ou effacé l'accord. Cela veut dire que dans une même œuvre, l'accord est parfois présent, parfois absent. Néanmoins, ces observations doivent être accompagnées d'une remarque importante : il est presque impossible d'établir un rapport de causalité entre l'absence de l'accord et le nombre des référents. De même, il paraît que l'insertion de l'article élidé ne joue aucun rôle lors de la détermination du référent du pronom *on*, l'insertion a plutôt une valeur stylistique.

De l'analyse résulte aussi que le référent du pronom *on* peut être interprété comme très précis (par exemple une personne concrète), comme généralement applicable (par exemple un proverbe) ou comme totalement imprécis (par exemple dans les textes où le

réfèrent du pronom *on* reste volontairement inconnu). Ces trois interprétations ont été regroupées afin de former une échelle de la détermination contenant trois degrés de détermination du pronom *on* : (1) explicitation, (2) vérité générale, (3) indétermination. Chaque degré est composé de plusieurs types d'équivalents. Il faut toutefois prendre en considération la différence entre l'appartenance à un degré de détermination en général et dans un contexte concret. Ainsi, chaque exemple a été classé individuellement. C'est pour cela que l'approche scalaire a été adoptée avec l'objectif de mieux saisir la nature transitoire de certaines valeurs référentielles du pronom *on*.

De nombreuses observations ont été effectuées. Mentionnons au moins quelques-unes. Dans le sous-corpus Core, la première et la troisième personne du pluriel sont les équivalents les plus utilisés en tchèque et en espagnol. De l'autre côté, l'équivalent *la gente/alguien* n'est presque pas présent alors que son homologue tchèque *člověk/někdo* est la troisième catégorie la plus souvent employée. Contrairement aux textes littéraires qui sont dotés d'une large variété d'équivalents, le passif réfléchi considérablement prédomine dans le sous-corpus Acquis dans les deux langues. En outre, le réfèrent indéterminé est plus répandu dans les textes juridiques que dans les textes littéraires.

L'étude des équivalents du pronom *on* en tchèque et en espagnol a prouvé que la détermination du réfèrent est un sujet extrêmement délicat et problématique. L'analyse de l'échantillon des 600 exemples (les textes littéraires) et des 1050 exemples (les textes juridiques) a démontré que les versions tchèque et espagnole ne se correspondent pas à cent pour cent. Au contraire, il est possible de déduire à partir de l'étude des textes littéraires que les équivalents ne se correspondent qu'en 40,5 % des cas. La tendance dans les textes juridiques est un peu plus convergente, le taux de correspondances atteint 50,1 %. Il faut toutefois avoir conscience du fait que l'hypothèse sur la similarité des équivalents a donc partiellement été confirmée puisque les types d'équivalents trouvés dans les textes se ressemblaient. Cependant, ces types d'équivalents n'étaient pas utilisés dans les mêmes contextes. L'hypothèse a donc été aussi partiellement infirmée.

Or, ce n'est pas seulement le type d'équivalent en soi qu'il faut prendre en compte, c'est aussi le degré de détermination auquel il appartient. Au niveau de ce degré, les versions tchèque et espagnole dans les textes littéraires se correspondent un peu plus – 54,3 % des équivalents se correspondent. Dans les textes juridiques, le taux de correspondance s'élève même à 60,8 %. Il est toutefois essentiel de regarder l'échelle

de la détermination en sa globalité. Ainsi, même s'il n'y a pas de question d'une fixation du degré de détermination objective, les différentes combinaisons des degrés de détermination en tchèque et en espagnol donnent à entendre que les traductions ne sont pas si différentes. En guise d'exemple : l'indétermination dans une langue et l'explicitation dans l'autre n'apparaissent presque jamais, les nuances s'effectuent plutôt entre l'indétermination et les vérités générales car les limites ne sont pas complètement nettes. Ainsi, l'hypothèse sur l'interprétation assez différente du référent du pronom *on* dans les traductions tchèques et espagnoles a été infirmée.

En outre, l'hypothèse que contrairement à celle des équivalents dans le sous-corpus Core, l'échelle de la détermination des équivalents dans le sous-corpus Acquis ne sera pas proportionnellement représentée, a été partiellement confirmée. L'indétermination a prédominé dans les deux langues (83,2 % des équivalents en tchèque, 92 % des équivalents en espagnol). Des traces du référent déterminé (tels que les pronoms personnels renvoyant à un référent concret ou les noms) peuvent toutefois être trouvées même dans les textes juridiques, bien qu'il y soit peu fréquent.

Finalement, la dernière hypothèse prévoyant que les référents du pronom *on* dans les textes juridiques seront placés plus bas sur l'échelle de la détermination a été confirmée. Néanmoins, contrairement aux attentes, il semble que les textes juridiques n'ont pas souvent recours à l'usage du degré des vérités générales, mais ils préfèrent le degré de l'indétermination totale.

Pour conclure, ce mémoire a dévoilé de nombreux facteurs qui théoriquement pourraient jouer un rôle lors de l'interprétation du référent du pronom *on*, mais qui d'habitude ne fonctionnent pas de cette manière : par exemple l'accord du verbe *être* avec l'adjectif ou le participe passé qui reste peu fréquent. L'analyse contrastive des équivalents tchèques et espagnols a montré que le plus grand impact dans cette problématique est celui du contexte linguistique (immédiat et large), mais aussi extralinguistique qui permet de découvrir le degré de détermination du pronom *on*.

Les échelles de la détermination regroupant les différents types d'équivalents du pronom *on* ont été construites et ces équivalents trouvés en tchèque et en espagnol y ont été placés. Ainsi, les tendances générales observées dans les corpus indiquent que le référent du pronom *on* demeure le plus souvent indéterminé dans les textes juridiques, tandis que

les textes littéraires ont le plus souvent recours au pronom *on* afin d'introduire des vérités générales. Pourtant, la prédominance d'un degré de détermination n'est évidente que dans les textes juridiques ; les textes littéraires sont plus homogènes (au sens que les degrés de détermination sont dans une certaine mesure également représentés).

Ainsi, il faut toujours regarder non seulement la phrase, mais aussi son contexte linguistique et extralinguistique étant donné que le type du texte (littéraire ou juridique) joue un rôle considérable. Néanmoins, il faut prendre en considération le fait que le contexte n'est pas toujours interprété de la même manière, même dans le cadre d'un même type de texte. Et c'est cela qui rend l'usage du pronom *on* si unique, son caractère multi-référentiel et son interprétation variable.

## 5 PISTES POUR DES RECHERCHES ULTÉRIEURES

Bien sûr, la présente étude n'est pas exhaustive car la problématique du fonctionnement du pronom *on* et de la détermination de son référent est très large. C'est pour cela qu'il est possible d'introduire quelques points de départ pour des recherches ultérieures.

Premièrement, le contexte a souvent été mentionné. Cependant, une étude plus centrée sur les propriétés pertinentes de ce contexte pourrait aller encore plus dans les profondeurs du sujet. Par exemple, les marqueurs spatio-temporels ont été remarqués, mais aucune attention spéciale n'y a été prêté. Aussi, une étude basée sur un corpus plus large (surtout dans le cas des textes littéraires) pourrait développer les constatations sur la nature et l'interprétation du pronom *on*. Voici donc une des possibilités du développement.

Deuxièmement, dans le cadre du travail, une brève comparaison des fréquences relatives du pronom *on* dans les textes littéraires a été présentée en s'appuyant sur les données du corpus référentiel Frantext. Or, une comparaison non seulement quantitative (les fréquences), mais aussi qualitative (l'accompagnement lexical ou les référents du pronom *on*) pourrait améliorer la force probante de la présente étude. Cependant, jusqu'à présent, le corpus Frantext disposait d'un moteur de recherche qui ne facilitait pas ce type de recherche. En juin 2018, la première version de Frantext a été remplacée par une nouvelle version 2. La nouvelle interface permet de faire des recherches dans des textes lemmatisés et étiquetés, il sera donc possible de mettre en relation la présente étude avec le corpus référentiel Frantext pour affiner et encore mieux interpréter les observations dans les travaux ultérieurs.

Finalement, une analyse des traductions au sens rétrograde, c'est-à-dire du tchèque et de l'espagnol en français, pourrait être effectuée. Il serait possible d'en dégager d'autres tendances et donc d'observer le fonctionnement des procédés d'impersonnalisation, par exemple si les équivalents du degré de l'indétermination se traduisent surtout par le pronom *on*, ou bien si il existe une autre traduction plus fréquemment utilisée.

## Listes des tableaux et figures

### Liste des tableaux :

Tableau 1 : Valeurs référentielles du pronom on selon le degré de détermination.....	15
Tableau 2 : Variété des équivalents du pronom on en espagnol .....	20
Tableau 3 : Variété des équivalents du pronom on en tchèque .....	24
Tableau 4 : Les verbes les plus fréquents accompagnant le pronom on .....	38
Tableau 5 : Composition des équivalents du pronom on en tchèque et en espagnol (dans les textes littéraires).....	41
Tableau 6 : Correspondance entre les équivalents – proportion de la totalité de l'échantillon (dans les textes littéraires).....	43
Tableau 7 : Correspondance des catégories d'équivalents en tchèque et en espagnol (dans les textes littéraires) .....	44
Tableau 8 : Composition des équivalents du pronom on en tchèque et en espagnol (dans les textes juridiques) .....	55
Tableau 9 : Correspondance entre les équivalents – proportion de la totalité de l'échantillon (dans les textes juridiques) .....	57
Tableau 10 : Correspondance des catégories d'équivalents en tchèque et en espagnol (dans les textes juridiques) .....	58

### Liste des figures :

Figure 1 : Distribution proportionnelle des lemmes dans les textes littéraires.....	36
Figure 2 : Distribution de fréquence des lemmes dans les textes juridiques.....	37
Figure 3 : L'échelle de la détermination (dans les textes littéraires).....	45
Figure 4 : L'échelle de la détermination (dans les textes juridiques) .....	60
Figure 5 : Comparaison des échelles de la détermination dans les textes littéraires .....	71
Figure 6 : Comparaison des échelles de la détermination dans les textes juridiques .....	71

## Listes des exemples analysés

(1) *J'étais titularisé, **on avait eu** la bonté de me donner un avenir et je le proclamais enchanteur mais, sournoisement, je l'abominais.* (Jean-Paul Sartre, *Les Mots*)

- *Obdržel jsem titul, **rodina** mi laskavě dala budoucnost, prohlašoval jsem, že je okouzlující, potměšile jsem ji však proklínal.* (traduit par Dagmar Steinová)

(2) *Le directeur lui avait dit qu'il comprenait très bien, mais qu'**on ne pouvait pas** faire d'exception, qu'il allait voir, mais qu'en somme la situation était grave et que **l'on ne pouvait rien** décider.* (Albert Camus, *La Peste*)

- *Ředitel mu odpověděl, že to vše dobře chápe, ale že **nemůže** dělat výjimky, že se mu vynasnaží pomoci, ale že situace je vážná a že **on sám nemůže** o ničem rozhodovat.* (traduit par Milena Tomášková)
- *El director le había respondido que lo comprendía muy bien, pero que **no podía** hacer excepciones, que vería, pero que, en suma, la situación era grave y que **no se podía** decidir nada.* (traduit par Rosa Chacel)

(3) *Grouillez-vous de boire ça, et puis **on va** danser.* (Boris Vian, *J'irai cracher sur vos tombes*)

- *Hod'te to do sebe a pak si **půjdeme** zatancovat.* (traduit par Martin Řízek et Petr Turek)
- *Terminatelo de prisa y **nos vamos** a bailar.* (traduit par Jordi Martí Garcés)

(4) *À les voir s'éparpiller, gesticuler minutieusement de la sorte et se perdre en dentelles de mouvements saccadés et follement inutiles, **on en demeurerait** découragé jusqu'au marasme.* (Louis-Ferdinand Céline, *Voyage au bout de la nuit*)

- *Když jsem je tak viděl běhat, přesně a svědomitě gestikulovat a mořit se s girlandami rázných a bláznivě zbytečných pohybů, **býval jsem** z toho na smrt sklíčený.* (traduit par Jaroslav Zaorálek)

(5) *Si **on va** doucement, on risque une insolation.* (Albert Camus, *L'Étranger*)

- *Když **jde člověk** pomalu, lehko si užene úpal.* (traduit par Miloslav Žilina)
- *Si **uno anda** despacio, corre el riesgo de una insolación.* (traduit par José Ángel Valente)



(6) Dans les cas de décès, **on a** toujours tendance à dire une connerie de ce genre. (Michel Houellebecq, *Les Particules élémentaires*)

- V případě úmrtí **máme** vždy sklon k podobným žvástům. (traduit par Alan Beguivin)

(7) Les Henrouille eux, n'en revenaient pas d'avoir passé à travers la vie rien que pour avoir une maison et comme des gens qu'**on vient de désemmurer** ça les étonnait. (Louis-Ferdinand Céline, *Voyage au bout de la nuit*)

- Henrouillovi nevyházeli z údivu nad tím, že prožili celý život jenom proto, aby měli vlastní dům; udivovalo je to jako lidi, co je zrovna **vytáhli** ze sklepa, kde byli zazděni. (traduit par Jaroslav Zaorálek)
- Los Henrouille, en cambio, no salían de su asombro por haber pasado por la vida nada más que para tener una casa e, igual que las personas a las que **acaban de sacar** de un encierro entre cuatro paredes, les resultaba extraño. (traduit par Carlos Manzano)

(8) Entre Corneille et Pardaillan, **on me sommait** de choisir. (Jean-Paul Sartre, *Les Mots*)

- **Byl jsem vyzván**, abych si vybral mezi Corneillem a Pardaillanem. (traduit par Dagmar Steinová)

(9) Ils veulent avoir lu le livre recommandé par leur club, celui dont **on parle**, et ils se moquent bien de ce qu'il y a dedans. (Boris Vian, *J'irai cracher sur vos tombes*)

- Lidé si chtějí hlavně přečíst knihu doporučenou jejich klubem, knihu, o níž **se mluví**, a kašlou na to, co je uvnitř. (traduit par Martin Řízek et Petr Turek)
- El libro que quieren leer es el que recomienda su club, el libro del que **se habla**, y su contenido les importa un bledo. (traduit par Jordi Martí Garcés)

(10) Dans le silence qui suivit cette démonstration vocale, **on entendit** nettement une mouche traverser l'atmosphère de la pièce avant de se poser sur le visage de Jane. (Michel Houellebecq, *Les Particules élémentaires*)

- V tichu následujícím po této hlasové ukázce **bylo** zřetelně **slyšet**, jak pokojem proletěla moucha a usedla na Janinu tvář. (traduit par Alan Beguivin)

(11) Quand le mystère est trop impressionnant, **on n'ose pas** désobéir. (Antoine de Saint-Exupéry, *Le Petit prince*)

- Cuando el misterio es demasiado impresionante, **es imposible** desobedecer. (traduit par Bonifacio del Carril)

(12) *Il y avait un coin d'où **on ne voyait pas** la route quand **on était assis**.* (Boris Vian, *J'irai cracher sur vos tombes*)

- *Bylo tam zákoutí, odkud **nebylo od země vidět** na silnici.* (traduit par Jaroslav Zaorálek)
- *Había un rincón desde el que **no se veía** la carretera **estando sentado**.* (traduit par Jordi Martí Garcés)

(13) ***On rappelle** à ce sujet l'avis exprimé par le Comité des régions sur la communication de la Commission sur "Accroître l'emploi des travailleurs âgés et différer la sortie du marché du travail" (adopté au cours de la 56ème session plénière, qui s'est tenue le 29 septembre 2004).*

- ***Výbor by chtěl odkázat** na své stanovisko (přijaté v průběhu plenárního zasedání dne 29. září 2004) ke sdělení Komise týkající se Zvyšování zaměstnanosti starších pracovníků a prodlužování odchodu z trhu práce.*

(14) *Territorialité : on transcende le concept de proximité territoriale et **on le remplace** par un concept de relation de réseau et de valorisation.*

- *Lokalita: koncepci zeměpisné blízkosti **nahrazuje** koncepce síťových vazeb a maximálního využívání různých kulturních identit partnerů.*

(15) *Dans le cadre des stratégies de développement rural appliquées à la méthode LEADER, **l'on agit** avec une grande efficacité sur le patrimoine rural, compris dans son sens le plus large.*

- ***Ženy** hrají významnou úlohu v diversifikaci venkovského hospodářství, často ale čelí specifickým překážkám, jako je omezený přístup ke kapitálu. Pozměňovací návrh je zaměřený tak, aby podpořil opatření k odstranění těchto překážek. V rámci strategií rozvoje venkova EAFRD úspěšně **podporují** venkovské dědictví v širším slova smyslu.*

(16) ***L'on peut voir** que cette approche de la Commission [...].*

- ***Budeme si moci všimnout**, že toto ustanovení Komise [...].*
- *Más adelante **observaremos** que esta postura de la Comisión [...].*

(17) *Le producteur ou le fabricant d'un produit alimentaire doit décider si **l'on peut consommer** ce produit tel quel [...].*

- *Producent nebo výrobce potravinářského produktu musí rozhodnout, zda **může být** produkt **konzumován** jako takový [...].*
- *El productor o el fabricante de un producto alimenticio debe decidir si el producto está listo para **ser consumido** como tal [...].*

(18) *À cette fin, **on entend** par « impuretés végétales » des matières végétales visibles [...].*

- *Pro účely tohoto odstavce **se** "rostlinnými nečistotami" **rozumějí** části rostlin rozeznatelné pouhým okem [...].*
- *A estos efectos, **se entiende** por « impurezas de origen vegetal » materias de origen vegetal perceptibles a simple vista [...].*

(19) *Aux fins du présent règlement, **on entend** par « carcasse de porc »*

- *A los fines del presente Reglamento, **se entenderá** por « canal de cerdo » [...].*

(20) *Si **l'on voulait** poursuivre la logique de l'argumentation de l'Allemagne, les créanciers privés devraient davantage renoncer à leurs créances que les créanciers publics parce que ces abandons de créances sont déductibles de leur base imposable.*

- *Z logického **rozvíjení** argumentace Německa vyplývá, že věřitelé ze soukromého sektoru by museli prominout vyšší dlužnou částku než věřitelé z veřejného sektoru, protože jejich prominutí není zahrnuto do daňového výměru.*

(21) *Toutefois, compte tenu des informations dont **on dispose**, il convient de limiter l'exposition à l'oestradiol [...].*

- *Vzhledem ke **stávajícím** informacím je zapotřebí omezit expozici organismu účinkům 17-beta-estradiolu [...].*
- *Sin embargo, a la luz de la información **existente** es conveniente limitar la exposición al estradiol-17  $\beta$  [...].*

(22) ***L'on devrait** tenir compte du fait que dans les nouveaux pays membres d'Europe centrale et orientale, l'infrastructure est en mauvais état [...].*

- ***Je třeba** mít na paměti, že v nových členských státech ve střední a východní Evropě je infrastruktura v horším stavu [...].*
- ***Hay que** tener en cuenta asimismo que las infraestructuras en los nuevos Estados miembros de Europa Central y Oriental están en peor estado [...].*

*(23) En ce qui concerne le volume des ventes des autres producteurs communautaires, **on observe** une diminution de 7,1 % en volume et de 7,8 % en parts de marché entre 2001 et la période d'enquête.*

- Pokud jde o objemy prodeje ostatních výrobců ve Společenství, Ø poklesly mezi rokem 2001 a obdobím šetření o 7,1 % a jejich podíl na trhu se ve stejném období snížil o 7,8 %.*
- Por lo que respecta a los volúmenes de venta de otros productores comunitarios, Ø estos han descendido un 7,1 % en volumen entre 2001 y el período de investigación y un 7,8 % en cuota de mercado durante el mismo período.*

## Références bibliographiques

1. Akrobou, A. E. (2011). *Manual de traducción : francés/castellano, castellano/francés*. Madrid : Visión Libros.
2. Alloa, H., De Torres, S. M. (2005). *Hacia una lingüística contrastiva : francés-español*. Córdoba, Argentina : Comunicarte.
3. Atlani, F. (1984). *On l'illusioniste*. In Grésillon, A., Lebrave, J.-L., Atlani, F., Danon-Boileau, L., Simonin, J. *La Langue au ras du textes*, p. 12–29. Villeneuve d'Ascq : Presses Universitaires du Septentrion.
4. Beneš, P. (1965). Le pronom *on* en français et ses équivalents en roumain. In *Etudes romanes de Brno*, I, p. 171–188.
5. Blanche-Benveniste, C. (2003). Le double jeu du pronom *on*. In P. Handermann et al. (Eds.), *La syntaxe raisonnée*. Bruxelles : De Boeck & Larcier, Éditions Duculot, p. 43–56.
6. Blanche-Benveniste, C. (2005). Structure et exploitation de la conjugaison des verbes en français contemporain. In *Le français aujourd'hui*, 1, n° 148, p. 75–87.
7. Bouhours, D. (1671). Les entretiens d'Ariste et d'Eugene. Nouvelle edition. Accessible sur : <http://gallica.bnf.fr/ark:/12148/bpt6k6294306x/> [consulté le 27 février 2018].
8. Boutet, J. (1986). La référence à la personne en français parlé : le cas de *on*. *Langage et société*, 38, p. 19–50.
9. Chevalier, J.-C., Blanche-Benveniste, C., Arrivé, M., Peytard, J. (1997). *Grammaire du français contemporain*. Paris : Larousse.
10. Colin, J.-P. (1999). *Dictionnaire des difficultés du français*. Paris : Robert.
11. Čechová, M., Krčmová, M., Minářová, E. (2008). *Současná stylistika*. Prague : Nakladatelství Lidové noviny.
12. Čermák, F. (2010). Obecný subjekt (objekt) a jeho lexikální vyjádření: anglické *one*, německé *man*, francouzské *on* a české *člověk*. In Čermák, F., Kocek, J. *Mnohojazyčný korpus InterCorp: Možnosti studia*, p. 182–189. Prague : Nakladatelství Lidové noviny.
13. Dubois, J. (1965). *Grammaire structurale du français: nom et pronom*. Paris: Larousse,
14. Dubois, J., Lagane, R., Niobey, G., Casalis, D., Calais, J., Meschonnic, H. (1971). *Dictionnaire du français contemporain: manuel et travaux pratiques pour l'enseignement de la langue française*. Paris : Librairie Larousse,
15. Dušková, L. (1999). Man-Sätze in Czech and in English. In L. Dušková, *Studies in the English Language. Part 2*. Prague : Karolinum, p. 41–86.

16. Fernández Ballón, M., Monnerie-Goarin, A. (1994). *Gramática esencial de francés*. Paris : Larousse.
17. Fournier, N. (2002). *Grammaire du français classique*. Paris : Belin.
18. Frei, H. (1929). *La grammaire des fautes*. Paris : Librairie Paul Geuthner.
19. Gjesdal, A. M. (2009). *Étude sémantique du pronom ON dans une perspective textuelle et contextuelle*. Accessible sur : <http://www.revue-texto.net/index.php?id=2417/> [consulté le 13 avril 2018].
20. Grepl, M., Karlík, P. (1998). *Skladba češtiny*. Olomouc : Votobia.
21. Grevisse, M., Goose, A. (2008). *Le bon usage: grammaire française : Grevisse langue française*. Bruxelles: De Boeck.
22. Guilbert, L. (1971) *Grand Larousse de la langue française: en six volumes*. Paris : Larousse.
23. Hendrich, J., Radina, O., Tláškal, J. (2001). *Francouzská mluvnice*. Plzeň : Fraus.
24. Landragin, F.; Tanguy, N. (2014). Référence et coréférence du pronom indéfini *on*. *Langages* 195, p. 99–115.
25. Narjoux, C. (2002). « *On*. Qui. *On* » ou des valeurs référentielles du pronom personnel indéfini dans *Les Voyageurs de l'Impériale* de Louis Aragon. In: *L'Information Grammaticale*, 92, p. 36–45.
26. Neumann, J., Hořejší, V., Vlasák, V., Kolář, V. (1992). *Velký francouzsko-český slovník. II. díl, L-Z*. Deuxième édition. Prague : Academia.
27. Obradović, R. (2013). Le pronom indéfini *on* en français contemporain et ses équivalents en serbe. In *ЗБОРНИК РАДОВА ФИЛОЗОФСКОГ ФАКУЛТЕТА*, 1, XLIII, p. 67–89.
28. Parquet, M. (2013). *Introduction générale au droit*. Breal : Lexifac Economie Droit.
29. Păuna, M. M. (2013). Le fonctionnement du pronom *on* en approche contrastive français-roumain. In *Bulletin scientifique en langues étrangères appliquées*, 1.
30. Radina, O. (1977). *Francouzština a čeština – systémové srovnání dvou jazyků*. Prague : Státní pedagogické nakladatelství.
31. Riegel, M., Rioul, R., Pellat, J.-C. (2009). *Grammaire méthodique du français*. 4e édition. Paris : Presses Universitaires de France.
32. Tláškal, J. (2005) K vyjadřování všeobecného lidského konatele (kontrastivní pohled na francouzské *on*). In Blatná, R., Petkevič, V. *Jazyky a jazykověda: sborník k 65. narozeninám prof. Františka Čermáka* (p. 375–394). Prague : Filozofická fakulta Univerzity Karlovy, Ústav Českého národního korpusu.

33. Wilmet, M. (2007). *Grammaire critique du français*. Paris : De Boeck.
34. Yao, K. (2016). Referencias y equivalencias del pronombre *on* francés en el español. *Revista de Lenguas Modernas*, 25, 115–130.
35. Zavadil, B., Čermák, P. (2010). *Mluvnice současné španělštiny: lingvisticky interpretační přístup*. Prague : Karolinum.

## Corpus consultés

1. FRANTEXT, version 1. Université de Lorraine: Analyse et traitement informatique de la langue française, 2018 [online]. [consulté le 15 janvier 2018]. Accessible sur : <http://www.frantext.fr/>
2. Nádvorníková, O. – Vavřín, M.: Corpus InterCorp – français, version 10 du 1<sup>er</sup> décembre 2017. Ústav Českého národního korpusu FF UK, Prague 2017 [online]. [consulté le 24 février 2018]. Accessible sur : <http://www.korpus.cz/>
3. Čermák, P. – Vavřín, M.: Corpus InterCorp – espagnol, version 10 du 1<sup>er</sup> décembre 2017. Ústav Českého národního korpusu FF UK, Prague 2017 [online]. [consulté le 24 février 2018]. Accessible sur : <http://www.korpus.cz/>
4. Rosen, A. – Vavřín, M. – Zasina, A.: Corpus InterCorp – tchèque, version 10 du 1<sup>er</sup> décembre 2017. Ústav Českého národního korpusu FF UK, Prague 2017 [online]. [consulté le 24 février 2018]. Accessible sur : <http://www.korpus.cz/>

## Résumé en français

Le présent mémoire de licence analyse le fonctionnement du pronom *on*, plus précisément la détermination de son référent. Etant donné que le référent de ce pronom demeure souvent vague, l'étude s'appuie sur ses équivalents en tchèque et en espagnol dans les textes littéraires et juridiques afin d'observer les différences entre son usage dans des contextes divers.

Le travail est divisé en trois parties : la partie théorique, méthodologique et empirique. D'abord, le pronom *on* est examiné du point de vue théorique, c'est-à-dire étymologique, syntaxique, stylistique et sémantique car il représente un défi à de nombreuses facettes. Les points de repère contrastifs démontrent les différentes perspectives référentielles sur lesquelles l'échelle de la détermination est construite.

Du point de vue méthodologique, l'analyse travaille avec les données issues des sous-corpus Acquis (pour les textes juridiques) et Core (pour les textes littéraires) du corpus InterCorp. La question de recherche principale est la suivante : Comment est-il possible d'identifier le degré de détermination du référent du pronom *on* en s'appuyant sur ses équivalents en tchèque et en espagnol ? Afin de répondre à cette question, de nombreuses questions de recherche partielles sont introduites.

Dans la partie empirique, c'est d'abord la fréquence relative des occurrences du pronom *on* qui est étudiée. L'hypothèse sur la prédominance du pronom *on* dans les textes littéraires est confirmée en s'appuyant sur la différence considérable entre la fréquence relative 804,35 ipm dans le sous-corpus Core et 306,03 ipm dans le sous-corpus Acquis. Puis, les données signalent que l'accord de l'adjectif et du participe passé avec le verbe n'est pas souvent employé, ce qui rend la détermination plus difficile. Pour ce qui est du voisinage du pronom *on*, le contexte immédiat dans les textes juridiques est influencé par la prédominance du verbe entendre, tandis qu'aucun verbe ne prédomine autant dans les textes littéraires dans lesquels une plus large diversité lexicale est percevable.

Ensuite, les types d'équivalents constitués à partir des données sont introduits. Dans les textes juridiques, le passif réfléchi est le procédé le plus utilisé, le tchèque toutefois offre une palette d'équivalents plus vaste que l'espagnol. Dans les textes littéraires, le pronom *on* est souvent traduit par la première ou la troisième personne du pluriel dans les deux langues.



Cependant, contrairement au tchèque, l'espagnol n'utilise pas beaucoup les équivalents *la gente*, *alguien* ou le passif promotionnel.

Au premier aspect, étant donné que les types d'équivalents ne se correspondent pas à cent pour cent, il semble que le référent n'est pas objectivement déterminable. Or, le regroupement des équivalents en trois degrés de détermination témoigne d'une autre tendance.

Les correspondances des degrés de détermination en tchèque et en espagnol constituent plus que la moitié des occurrences dans les textes littéraires et juridiques. En outre, dans les textes juridiques, une forte tendance d'uniformité est observable car la majorité absolue des référents du pronom *on* sont des référents indéterminés. De l'autre côté, les référents du pronom *on* dans les textes littéraires sont plus proportionnellement représentés dans les trois degrés de détermination.

Ainsi, il en résulte que les éléments clé pour l'observation du degré de détermination sont les suivants. Premièrement, c'est le contexte immédiat du pronom *on* qui influence le degré de détermination, par exemple l'accord de l'adjectif ou du participe passé avec le verbe, ou le voisinage verbal. Deuxièmement, c'est le contexte plus large dans lequel peuvent apparaître des notions qui aident à déterminer le référent. Troisièmement, c'est le contexte extralinguistique, c'est-à-dire la nature du texte, l'usage, l'objectif du texte et d'autres aspects.

Finalement, il faut aussi prendre en compte les contraintes rencontrées. D'abord, c'est l'étendue du corpus limitée et la composition spécifique des textes qui peuvent influencer les résultats. Puis, l'alignement et l'étiquetage imprécis du corpus (surtout du sous-corpus Acquis) ainsi que les mélanges des variantes linguistiques rendent la recherche plus difficile.

En ce qui concerne les recommandations pour des recherches ultérieures, il conviendrait approfondir l'analyse en examinant le contexte large d'une manière plus détaillée. Il est à noter qu'une autre piste constitue la comparaison des résultats issus d'InterCorp avec des corpus référentiels (tels que Frantext). Enfin, c'est aussi une analyse des traductions au sens rétrograde qui permettrait de perfectionner l'étude des référents du pronom *on*.

## Résumé v češtině

Tato bakalářská práce si klade za cíl analyzovat fungování zájmena *on*, přesněji řečeno míry determinace jeho reference. Jelikož je jeho reference často vágní, tato studie pracuje s českými a španělským ekvivalenty tohoto zájmena v beletristických a právnických textech, díky čemuž lze pozorovat rozdíly mezi jeho použitím v různých kontextech.

Práce je rozčleněna do tří částí: teoretické, metodologické a empirické. V teoretické části je zájmeno *on* zkoumáno z několika úhlů pohledu, a to z pohledu etymologie, syntaxe, stylistiky a sémantiky. Dále jsou představena kontrastivní východiska překladových ekvivalentů tohoto zájmena v češtině a španělštině, na nichž staví škála determinace.

Z metodologického pohledu se jedná o korpusovou analýzu, která se opírá o subkorpusy *Acquis* (právníké texty) a *Core* (beletristické texty) paralelního korpusu *InterCorp*. Práce si klade za cíl najít odpověď na základní výzkumnou otázku: Jakým způsobem lze určit stupeň determinace zájmena *on* za pomoci jeho českých a španělských ekvivalentů? Kromě této hlavní otázky je v práci zodpovězeno i několik dílčích otázek, které se zabývají různými kvantitativními i kvalitativními aspekty zkoumání zájmena *on*.

V empirické části jsou nejprve porovnány frekvence výskytů tohoto zájmena v beletrii a textech Evropské unie. Je potvrzena hypotéza o jeho vyšší relativní frekvenci v prvním typu textů, o čemž svědčí hodnoty 804,35 ipm v subkorpusu *Core* a 306,03 ipm v subkorpusu *Acquis*. Z následující části je patrné, že ke shodě přídavného jména nebo přičestí minulého se slovesem dochází zřídka, čímž je determinace ztížena. Z pohledu lexikálního obsazení lze konstatovat, že bezprostřední kontext ovlivňuje v právních textech zejména sloveso *entendre*, zatímco literárním textům žádné sloveso takto nedominuje, panuje v nich větší diverzita sloves.

To má vliv také na jednotvárnost překladových ekvivalentů. V případě právních textů je nejčastějším prostředkem vyjádření zájmena *on* zvrtné pasivum, nicméně v češtině je patrná širší škála překladových protějšků než ve španělštině. V beletrii se v obou jazycích nejčastěji používá k překladu zájmena *on* první a třetí osoba plurálu. Na rozdíl od češtiny se však španělština příliš často neuchyluje k použití neurčitých zájmen *la gente* a *alguien* či opisného pasiva.

Vzhledem k tomu, že si typy překladových ekvivalentů stoprocentně neodpovídají, může se na první pohled zdát, že referenci objektivně určit nelze. Tyto typy je ovšem nutné

seskupit do větších celků, jejichž ukazatelé budou mít větší výpovědní hodnotu, a to do tří stupňů determinace.

Ke shodě ve stupni determinace v češtině a španělštině dochází ve více než polovině pozorovaných příkladů jak v beletrii, tak v právních textech. Legislativa se navíc v naprosté většině případů uchyluje k nejnižšímu stupni determinace. Tento typ textů je, co se neurčitosti referenta týče, značně stereotypní. Naproti tomu jsou stupně determinace reference zájmena *on* v beletristických textech rovnoměrněji zastoupeny.

Z toho vyplývá, že klíčovými prvky pro určení stupně determinace jsou následující aspekty. Zaprvé jde o bezprostřední kontext zájmena *on*, například shoda adjektiva nebo minulého přičestí se slovesem či lexikální obsazení slovesa. Zadruhé je třeba věnovat si širšího kontextu, v němž se mohou objevit další indikátory determinace. Zatřetí je nutné zohlednit i extralingvistický kontext, tedy typ textu, úzus, jeho cíl a další.

Nakonec je zapotřebí zmínit i úskalí výzkumu. Jmenujme například vliv omezeného rozsahu korpusu a jeho specifického složení. Výzkum ztěžuje i mísení jazykových variant, nepřesné zarovnání a tagování korpusu (především subkorpusu *Acquis*). Co se týče doporučení pro další výzkum, bylo by vhodné analýzu rozšířit právě o pozorování širšího kontextu, především v případě beletristických textů. K dalším možnostem rozšíření výzkumu patří srovnání s referenčními korpusy francouzštiny (například *Frantextem*) či ověření výsledků zpětným překladem.